RJP

The Richmond Journal of **Philosophy**

Issue **Seven** **Summer 2004**

In this issue...

Alan Thomas
on liberalism

Peter Goldie
on personality

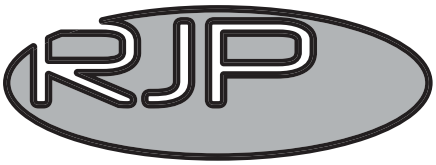
Garrath Williams
on moral responsibility

Christopher Hamilton
on sexual desire

Mat Carmody
on vagueness

Jennifer Booth
on scientific knowledge

Richmond upon Thames College



The Richmond Journal of Philosophy

Issue seven
Summer 2004

Philosophy Department

Richmond upon Thames College

Egerton Road

Twickenham

Middlesex

TW2 7SJ

United Kingdom

email: rjp@rutc.ac.uk

www.rutc.ac.uk/rjp

Contents

[Editorial]

Welcome to the seventh issue of the Richmond Journal of Philosophy.

We begin with a discussion by Alan Thomas on the nature of liberalism and then move onto Peter Goldie's consideration of the relationship between personality and the possibility of explaining how people can surprise us by what they do. This is followed by the second part of Garrath Williams' discussion of our understanding of moral responsibility. Having examined the views of Aristotle in the previous issue attention now turns to the views of Immanuel Kant. Our fourth paper by Christopher Hamilton focuses on sex, and in particular on the conception of sexual desire. After sex Mat Carmody invites us to think about vagueness, outlining the problems vagueness poses for our understanding and knowledge of the world and some possible responses. Scientific knowledge is often regarded as enjoying a special status. In our final paper Jennifer Booth reflects upon rival models of scientific knowledge in assessing how we should regard the nature of such knowledge.

Purpose of the Journal

The motivation for and ambition of the journal is to provide serious philosophy for students who are at an early stage in their philosophical studies. The style and content of the papers will be accessible to students who have yet to become hardened to the more technical and specialised journals of professional philosophy.

What do we mean by 'serious' philosophy? First, the content of the journal is not constrained by a remit to appeal to or reach the interested general public. Whilst the papers must speak to the needs of students who are relatively inexperienced in philosophy, they presuppose that their audience is actively engaged in philosophy. Second, the content is serious in its focus on the central areas of philosophy. One must beware of the dangers of trying to impose more precision on a subject than its nature will allow. Therefore, some degree of caution is called for in talking of the central areas of philosophy. Nonetheless, the big or traditional questions of metaphysics, epistemology, and ethics will provide the journal's centre of gravity. The third way in which the philosophy is serious is through the scope, variety and depth of analysis that can be achieved by the accumulation of papers over time. Moreover, each paper is not simply an introduction to one of the main topics on A-level, IB or degree courses. Such papers will indeed have a role in the journal, but they will not be the only kind. Our contributors will be offering original papers based on their own research. The journal will be a forum for the kind of critical engagement and debate that characterise the practice of philosophy. The fourth way in which the philosophy is serious is in the contributors themselves. The vast bulk of the papers will be written by professional philosophers engaged in both research and teaching.

About the [Editorial] Board

Stephen Grant is a full-time lecturer in philosophy at Richmond upon Thames College. He has also taught at King's College London where he is completing his doctorate on the emotions. His main interests are in the emotions, ethics and political philosophy. He has published on the ontological argument.

Dr Paul Sheehy teaches philosophy at Richmond upon Thames College and King's College London. His main areas of interest are in metaphysics, political and moral philosophy and the philosophy of the social sciences. His doctoral thesis was on the ontological and moral status of social groups, and he has published papers on social groups, voting and explanation and realism.

Paul Sperring is head of the philosophy department at Richmond upon Thames College and an A-level examiner in philosophy. He completed his undergraduate and masters studies at The University of Warwick, studying both analytic and continental philosophy. He is currently working towards his PhD at Birkbeck College. His research interests are metaphysics and the philosophy of mind.

[Editorial]

Alan Thomas

What does a [Liberal] Society demand of its Citizens?

There are many different forms of political organization in the world. What is it that makes a form of government for a particular society a liberal one? Within such a society, what demands does liberalism place upon the citizens who live within its boundaries? As an approximate answer to the first question, necessarily generalizing over centuries of liberal practice and theory, a liberal society is a society that prioritises, above other political values, the freedom of the individual as a political ideal. The way in which liberalism secures this priority is twofold: first, an individual is assigned certain basic rights and liberties. Second, to make these rights and liberties practical and politically effective, such a society also pursues an ideal of equality. As an approximate answer to the second question, the question that will be the main focus of this article, liberalism demands that its citizens be, well, liberal citizens. But there are more, or less, ambitious accounts of that in which liberal citizenship consists.

The balance between the two priorities within liberalism, between basic rights and equality, determines the different forms of liberal political theory. At one extreme, a liberal may assume that our basic moral rights are prior to politics and severely constrain the political process. The pursuit of equality, as a goal of a state's redistributive policies, would represent an interference with certain basic moral rights, such as the right to acquire, hold and transfer private property. This can be represented as the view that, in the conflict between freedom and equality, freedom should be given a stronger emphasis. More egalitarian liberalisms see the reality and efficacy of certain basic rights as themselves underwritten by a more demanding ideal of equality, perhaps one that only permits inequality in so far as doing so would improve the situation of the very worst off. On this view permitted inequalities must benefit the worst off so that, metaphorically, as the 'ceiling' of the rewards of the most highly paid rises, so it drags up the 'floor' with it, improving the lot of the comparatively worst off. This second view can be understood as the claim that equality is prior to freedom or, perhaps more accurately, as the view that a proper understanding of that which freedom consists in undermines any general

contrast between these two values. Equality, on this second view, is an essential condition of the meaningful exercise of those basic liberties protected by our most fundamental rights: perhaps those rights given protection by the constitution of the state. There is a very basic disagreement between these rival views about the relative importance of morality and politics, and which is allowed to constrain the other.

However, recent arguments within liberal political theory have focused not so much on this perennially controversial issue of the nature of the demands of equality, but on the wider question of the legitimacy of a liberal political order itself. If the debate between a more libertarian, freedom based and a more egalitarian, equality based liberal theory is about how we implement a liberal political theory, the issue of legitimacy asks why we should be liberals at all. The issue of legitimacy concerns the most fundamental of political issues, namely, what makes a particular form of government legitimate for the people that it governs. This is not the question of whether a citizen of a state agrees, in any particular case, with the policies of those that govern her. Accepting that a government is legitimate is compatible with very substantial disagreement about government policy. Indeed if, as many have argued, politics is fundamentally about accommodating disagreement

and preventing it spilling over into factionalism and violence, then this aspect of legitimacy is crucial to understanding what it is to set up a political order. When we collectively agree to be governed by a particular form of political organization and, in particular, to give it a monopoly over the legitimate use of force and to allow it to determine what is to happen rightly, even if that does not always co-incide with what we personally take to be right, we set up an order to govern our political life together that is legitimate. We accept, as it were, the fundamental ground rules of a game, not the particular moves that occur within it.

All liberal political theories claim that liberalism has certain fundamental advantages over other forms of political organization when it comes to the issue of legitimation. Liberal political theories claim that their legitimacy can be demonstrated to each citizen, on a rational basis. The reasons given for the legitimacy of a liberal political order can, moreover, be the same reason for each citizen or group of citizens: you don't have to give different reasons to different groups or different people. Furthermore, this process of justification can be transparent to all and does not rest on a lack of truthfulness or on coercion or deception. The legitimating narratives of other forms of political organization, the liberal claims, will be found wanting in comparison with these features of liberal legitimacy. Despotic regimes go through a sham process of legitimation, but its fraudulent character is brought out by the fact that the only thing sustaining that claim to legitimacy is the very same blatant exercise of power that the justification seeks to endorse.

All liberals seek to advance a legitimate political conception, but the more demanding a liberalism, the more it will try to demonstrate that it can be assented to by everyone. At this point the rarefied atmosphere of political theory makes contact with the practicalities of everyday politics, as it is undoubtedly true that the more egalitarian form of liberalism, that which justifies inequality solely in cases where doing so would benefit the worst off, bears less of a relation to the dominant political theories that underlie political practice in the contemporary United Kingdom and United States. It seems that a significant minority of the electorate of those democracies feel that a more demanding forms of liberal equality is too demanding: that it generates a system in which the State is too involved in the redistribution of resources that those citizens believe to be rightfully theirs. I note this thought without endorsing it; on a more plausible view, nothing is anyone's until the entire system of allocating and transferring resources for a whole society has been set up.

That is one ground on which one might be reluctant to embrace liberalism, even if one were to regard the view rejected as a legitimate option. But another aspect of the problem facing this particular form of egalitarian liberalism, a form associated with the American political philosopher John Rawls, strikes more fundamentally at the question of the legitimacy of the view itself. Suppose one came to think that the view was, in a sense, exclusionary, or sectarian? This is not the worry that, if one were subject to the demands of a Rawlsian society, one would take its requirements to be too demanding. This is the deeper anxiety that one could not be subject to such demands or, at least, that a person able to become subject to such demands would not be recognizably you as you would have given up on a valuable form of life that is very important to your moral identity. On what grounds might one think that?

On the grounds that liberalism represents a form in which we might choose to organize our social life, but not the only available form and that there are valuable ways of living, which would be incompatible with a liberal way of living. Clearly, liberalism has an individualistic bias, reflected in its starting from the individualistic notion of a right, or of a basic liberty. While liberalism in its different forms does have an account of the good of community, such goods have to emerge from within a context shaped by the rights and protections assigned to individuals. There are comprehensible, indeed, familiar ethical outlooks where individualism does not receive such an unqualified endorsement: ethical outlooks which emphasize the goods of a traditional outlook, or of a traditional way of life, perhaps involving certain taken for granted and hierarchical ways of organizing social life. From such a perspective organizing one's social life in terms of rights encourages strident self-assertion, adversarial social relations and a decline, not a reinforcement in sociability. Yet the liberal account of legitimacy clearly aims to bring everyone on board and not, in this way, to offer justifications of liberalism that appeal solely to those antecedently persuaded of its attractions. Rawls came to the view that his own early work, in which he developed a compelling case for a demanding liberal egalitarianism, had failed to be sufficiently attentive to this wider problem of legitimacy.

But was he right to be concerned in this way? Rawls himself accepted the point that no political view, not even liberalism, can accommodate every possible valuable way which humans can devise to live together. It is tempting to view liberalism as a maximally accommodating view, reflected in how broadly it understands the scope of its own legitimacy. But it would be a mistake to see liberals as trying to stock their own society with some maximally inclusive smorgasbord of options, a political analogue of the intergalactic bar in Star Wars (where aliens from many worlds share a drink, not a collective political life). Furthermore, liberals as a whole stand by the attractiveness of the basis of liberal legitimacy as opposed to those traditional legitimating narratives used to justify the divine right of kings to rule, or the authority of a traditional 'ruling class'. Those narratives rest on falsehood, or coercion, or both.

Those most untroubled by this issue are a group known as ethical liberals. Represented by the contemporary political philosophers William Galston and Stephen Macedo, the ethical liberal argues that it is both undeniably true, yet not to be regretted, that liberal political societies shape the dispositions and behaviour of those citizens governed by that very same liberal order. Corresponding to liberalism as a political theory there is liberalism as a personal ethical outlook. The two stand in a mutually supporting relationship. A liberal virtue ethic sees a person's ethical outlook as made up of certain typical virtues of character: liberal citizens are tolerant, open-minded, experimental in attitude and responsible. Responsibility is a key virtue. That is one explanation as to how liberal societies can afford to give their citizens a substantial sphere of life which is outside of the direct control of the state, and yet not collapse from the massive social costs which might be brought about if everyone within that state chose to behave in that sphere completely irresponsibly. A citizenry whose recklessness and lack of moderation led a majority to severe and chronic self-inflicted ill health, or that spent its time in the workplace either drunk or stoned, would very quickly bring the public services offered by a state to breaking point.

For ethical liberals, then, liberal political practice is matched by a liberal virtue ethic. Rawls came, in his later work, to be concerned that his account of liberal egalitarianism was

sectarian in the sense that it rested on too narrow a basis of justification: that it would appeal solely to those who were already committed to a liberal form of ethical life. Galston and Macedo accept this circularity in the justifications available to the liberal very clear-sightedly. Liberalism, like any other political view, will shape its citizens in its own image and that is not to be regretted. Macedo has even claimed that, through time, all liberal societies will converge on a single basic pattern and, more provocatively, that the best contemporary example of the pattern is contemporary California.

To articulate this dispute more clearly, take an example from the American political philosopher John Tomasi and divide people into three broad classes. The first group are, in their personal attitudes and behaviour, committed to the very aspects of character that Galston and Macedo call the liberal virtues. Open-minded, tolerant and experimental, this group of people value diversity and change in their personal and social lives, welcome exposure to novelty and difference and value crafting a distinctive and individual moral identity. The second group of people structure their lives by some traditional source of identity which is 'foundational' for them, whether religious or not.

This group dislike individual self-assertion and while they value the option of appealing to the individualistic rights discourse of liberalism, they do not value exercising that option as it seems to them a last resort that marks a breakdown in social relations. Their overall view is marked by respectfulness towards traditional sources of authority. The third group simply don't have a consistent and fully worked out set of principles or attitudes; a mixture of groups one and two, different parts of their lives involve different degrees of respect for tradition or authority, experimentation and 'standing out from the crowd'.

We can now put Rawls's concern this way: the first group will clearly have no problem accepting the legitimacy of liberalism not, at least, in terms of its justifiability to them from their own perspective. The third group are, shall we say, ripe for conversion: it will be possible so to mould their attitudes and expectations that they will come, through time, to resemble the first group. The point of concern is the second group. Do we, at this point, simply remind ourselves that liberalism cannot please everyone, that it does not aim to accommodate every possible valuable form of living together and that no-one is entitled to a level playing field? Through time, particularly as children born within this group are educated in the public education system and exposed to the wider influences of society, this group will find it more and more difficult to reproduce itself and will, slowly, 'go out of business'.

Is this to be regretted? I think Galston and Macedo are committed to the view that while this would, indeed, be regrettable, it is not regrettable in such a way that a liberal state itself is to blame for allowing it to happen.

Rawls's answer is that we should try and do more: that we ought to affirm unequivocally the importance of politics and its priority over morality. Politics imposes a special burden on citizens, which is that, when they come together to discuss fundamental matters of political legitimacy, we owe our fellow citizens a duty of restraint. What we restrain ourselves from is trying to justify legislation on the basis of our own comprehensive moral vision of the world, especially if we believe that this view is not widely shared. Reasonable people know that people can, blamelessly, disagree over fundamental issues about morality even if all the views under consideration are themselves reasonable (no-one cares if unreasonable world views go out of business - the passing of the Klu Klux Klan from liberal society is not regrettable from anyone's perspective). What Rawls tries to do is disjoin, or separate, our political commitment to liberalism from any particular general moral view that might support it. But this is not to detach politics from morality: the key words are 'any particular' view that might support it. What Rawls wants is a situation in which each particular general moral view offers its own proprietary support for liberalism from within its own particular perspective.

What you affirm is the common ground between you and your fellow citizens; the basis on which you affirm is that which you could not put forward to your fellow citizens without violating your duty of restraint. You know that if you put those considerations forward to others they would reject them as unreasonable for that purpose - of justifying political fundamentals - and they would be right to do so. Rawls called his later view of liberal egalitarianism 'modular' and I think what he meant by that metaphor is now clear. It is a module that 'slots in' to different general moral views of the world and is motivated from within each such reasonable view, while not being exclusively attached to any of them. It derives its power to motivate from its affirmation from within each of these particular general moral views.

Does this solve the problem? Rawls's aim is to give an account of the legitimacy of liberalism to all three of our representative groups. The second group are clearly the crucial test case. Can they affirm, from within their traditional worldview, liberalism construed as a political solution to a political problem: the problem of how to live together in a political community that we genuinely share, given that all sides agree that there can be blameless disagreement about the fundamentals of morality (as opposed to politics)? Perhaps equally importantly, can they affirm some adaptation or extension of their traditional worldview, adjusted to the particular conditions of a liberal society? Rawls believed that the answer to that question was 'yes'; even bearing in mind that liberalism does not attempt to accommodate all valuable forms of living together, it can go a little further in drawing a wider class of ethical outlooks within its ambit.

I have taken the issue of legitimacy as the focus of this article and there are now two clearly conflicting answers to the question of what a liberal society demands from its citizens for consideration. Ethical liberals ask that the citizens of a liberal society adopt the liberal virtues of criticism, open mindedness, a taste for experimentation and tolerance. Rawls, in his later work, asks for a different commitment, a demanding political ideal that is expressed in his duty of self-restraint. You are going to have to accept, when you put forward a political argument about constitutional fundamentals, that a certain kind of full or complete justification is unavailable to you and that you have, as it were, to 'stay on the surface' and affirm that which you share with all of your fellow citizens from within a general moral outlook on the world that, ex hypothesi, you do not share with them in the same way. But while in that respect you 'stay on the surface' the commitment that Rawls asks of you is by no means itself superficial, but clearly in its own way very demanding. Like his ethical liberal rivals, Rawls places the demands of citizenship centre stage, but he understood liberal citizenship as a certain narrowly defined political role. But given the priority and importance of politics it is the most important role that we can be called upon to take up.

Those unconvinced by Rawls's arguments fear that the same problem recurs: is it not true that we only affirm liberalism on the basis of those views that Rawls allows are reasonable general moral views? As I have noted, no one is concerned if, within a liberal society, the 'traditional forms of life' of the racist go out of business. But now is Rawls's restriction to reasonable views going to introduce a new circularity into what he is prepared to see justified within liberalism? Only reasonable views are allowed to count, and what handle do we have on reasonableness independently of these ways of living proving compatible with Rawls's later liberalism? I don't think that this criticism is justified, and the comparison with the 'ethical liberalism' of Galston and Macedo shows why.

The ethical liberal points to the legitimation of liberalism and says: liberalism recommends itself on its own merits. It meets a high standard of individual justification as legitimate without falsehood or coercion. In that sense any justification of it is circular: it rests on its own merits, but to what else can one appeal? The circularity of resting on one's own merits is still clearly different from the circularity of the 'justifications' offered by a despotic regime, where justifications are solely supported by the very exercise of power that the account tries to legitimate. Rawls accepts a degree of inevitable circularity in good justifications, those that do not rest on the brute causality of power, but, as it were, widen the circle. Our crucial second group, the test case for the transition within Rawls's views, can be brought within the terms of the liberal settlement whereas within ethical liberalism they cannot. That is an incremental gain, but a gain nonetheless, in arguments over the appeal of liberalism as legitimate means of organizing our political life together.

University of Kent.

Peter Goldie

What People will do: [Personality] and Prediction

I

People sometimes surprise us by the things that they do. What is the explanation of this phenomenon? This is a very general question, and I want here to treat it as such. I do, however, want to restrict the question to those surprising things that people do that are intentional actions, or things done for a reason in the particular sense that Elisabeth Anscombe was searching for in her book *Intention*.[1] For people sometimes do surprising things other than intentional actions. Talleyrand, the great diplomat who served the Ancien Régime, then the Revolutionary government, then Napoleon I, then the restoration of the monarchy with Louis XVIII, and then after the 1830 revolution with Louis-Philippe, was said never to do anything without intention, without a good reason. When he finally died in 1838, Metternich, surprised, made this famous remark: 'I wonder what he meant by doing that'.

People don't die intentionally, for reasons in the sense required for meaningful intentional action (suicide is something quite different). And that's what makes Metternich's remark into a nice joke, as well as indicating what is supposed to be special about action as such, according to which dying is not a kind of action.

So let me from now on limit my question to intentional action, which I will characterise as something done for a reason, consisting of a belief and an attitude towards this kind of action (let's from now on call it a desire, in the knowledge that this term is desperately vague; but nothing hangs on it here). And this belief and desire will give the answer why, in the sense we want, people do the things that they do; the belief and the desire rationalise the action.[2] Someone opens the door of the fridge, and he does this because he wants a beer and believes that opening the door is the best means of getting a beer, believing, as he does, that there is a beer in the fridge. Someone takes an umbrella on her walk, and she does this because she believes it's going to rain later on, because she wants not to get wet, and because she believes that taking an umbrella is the best means of avoiding getting wet. This kind of means-end, belief-desire explanation, it is said, can be extended to explain much more complex examples of intentional action.

I accept that, in respect of all intentional actions, these kinds of belief-desire explanations are available. Indeed, it seems to me that, in our thinking about action, our own and other people's, we just *take it for granted*-it is, I dare say, a priori-that belief-desire explanations are always going to be available; although, of course, they may not always be found.

However, a moment's reflection on these belief-desire explanations should reveal just how thin they are, partly because, in turn, they rely on such a thin notion of rationality. To rationalise an action in this thin sense is just to show how it could make sense for someone to do such a thing. And it can often make sense to do one of a diverse range of possible things. In a restaurant you are brought the wrong flavour of ice-cream: you ordered chocolate and you are brought strawberry. It would make sense to tell the waiter he has made a mistake; it would make sense to eat what you are given; it would make sense to leave it; it would make sense to throw the ice-cream on the floor and walk out of the restaurant; it would make sense to offer the ice-cream to your partner; it would make sense to pour the ice-cream onto your partner's lap; and so on. If you do one of these things, it's very likely that you didn't at the time consider all the other possible courses of action that were open to you, or perhaps any of them; the other possibilities might not have entered your mind. So we must not forget that a belief-desire explanation doesn't explain why someone did one thing *rather than another which was also open to him.*

The point emerges most starkly when we turn to prediction: I ask you to predict what a rational person (that's all we know about him or her) will do if he or she is brought the wrong flavour of ice-cream by a waiter in a restaurant. 'Well,' you will sensibly reply, 'it depends.'

One of the things it depends on is personality. Because I know my wife's personality, I can pretty reliably, I think, predict what she would do in such circumstances. However, I acknowledge that my prediction might be wrong: she might on a particular occasion surprise me and do something other than what would be expressive of her personality. So what she does must depend on more than just her personality.

To explain why I am right to acknowledge that my wife might surprise me in what she does, I need to turn to E. M. Forster's distinction, which he makes in his *Aspects of the Novel,* between flat and round characters. 'Flat characters', he says, 'were called "humours" in the seventeenth century, and are sometimes called types, and sometimes caricatures. In their purest form, they are constructed round a single idea or quality ... The really flat character,' he continues, 'can be expressed in one sentence such as "I will never desert Mr Micawber". There is Mrs Micawber-she says she won't desert Mr Micawber; she doesn't, and there she is'.[3] Flat characters are contrasted with round characters. 'The test of a round character,' Forster says, 'is whether it is capable of surprising in a convincing way. If it never surprises, it is flat. If it does not convince, it is flat pretending to be round. It [the round character] has the incalculability of life about it'.[4]

It is because flat characters have this 'single idea or quality' that they never surprise us; everything they do is determined by whatever quality or characteristic they are deemed to have, and they never act contrary to, or against, their type. We real life human beings, in contrast, like my wife, are round characters-all of us. So, by definition, we round characters are capable of surprising, and when we surprise, we do so in a convincing way.

The fact that people can surprise us by what they do shows that we expected them to do one particular thing, or one of a range of things, and in fact they did something different. Expectation in this context is, I think, something less that prediction, although it can extend to prediction. Expectation can involve simply taking some action or other for granted, rather as you expect-take it for granted-that someone who buys a cup of coffee in Caffé Nero will drink it and not throw it on the floor, or that a tennis player in the middle of a game will try to return the ball and not try to imitate Marlon Brando. Expecting what someone will do in some circumstance or other, as my ice-cream example shows, and as these examples show, involves more than just expecting what any rational person would do: the idea of rationality has to be thickened out with *personality.* So, given a particular person's personality, my wife's for example, we can expect that she will do one kind of action amongst a range of possible rational actions, and expect that she will not do other kinds of action which it would be equally rational to do. For someone else with a different personality from my wife's, we may expect them to do something else, and not what we expect my wife to do. This, personality and expectation or prediction, is the first thing I will try to explain.

The second thing I will try to explain is the confounding of our expectations and predictions-the fact that we can be surprised, in a convincing way, when people do something other than what we expect-when they act contrary to what their personality traits lead us to expect.

Then, thirdly, I will try to draw some practical conclusions from all this-things that we round characters might bear in mind in leading our lives, and in judging ourselves and others. But let me begin by saying something about what I mean by personality.

II

What is a personality trait? Personality traits-and here I include character traits in this category-are dispositions. All sorts of things have dispositions. This glass is fragile. Its being fragile (its fragility) is a disposition of the glass. Dispositions like these can be understood in terms of 'if-then' conditional statements: 'if this glass were to be thrown against the wall, then it would break'.

When we come to personality traits, to say what the 'if-then' conditional looks like *in general* is, I think, a hopeless task, because personality traits are so disparate in kind. There are ways of acting, such as being charming; there are relatively enduring temperaments, such as being cheerful, being nervous and being gloomy; there are relatively enduring emotional dispositions, such as irascibility and being envious; there are relatively enduring preferences and values, such as being a book-lover, being a foodie, liking football, and disliking authority figures; there are talents, such as being quick-witted; and there are character traits, which are, roughly, relatively enduring dispositions to have certain kinds of motives in certain kinds of situation, and thus to act in certain kinds of ways.

'If-then' conditional statements for personality traits are what Gilbert Ryle has called 'inference-tickets'-'season tickets' which allow us to explain and predict motive and action on a particular occasion.[5] For example, knowing that someone is a polite and considerate person allows us to predict that she will have particular polite and considerate motives if she sees that her fellow dinner-guest needs the salt, and thus that she will pass the salt in a polite way.

Flat characters have a single personality trait, which can be utterly relied upon to issue in the relevant kind of action on all occasions. Thus flat characters do not-they cannot-surprise us. But, as Forster rightly says, flat characters are only to be found in novels; it is a kind of prejudice-a kind of stereotyping-to flatten out real life people. So, for us real life people, we need to build into our notion of personality this capability of surprising us. Any notion of reliability, of predictability, in motive and action must allow for the possibility of a failure to be motivated or to act as our ascription of personality traits allows us to infer. One might put it like this: an inference-ticket, a season ticket, can fail on an occasion to get you to your expected destination.

We could just build in to the 'if-then' conditional a 'normally', a 'usually', or an 'all other things being equal', and this would certainly allow for the possibility of being surprised: a polite person will *normally* or *usually* have polite motives and act politely when it is appropriate. Fair enough. However, although this might explain how we can be surprised-on those occasions when the 'normal' or 'usual' doesn't apply-it isn't going to be enough to explain how we can be surprised *in a convincing way.* So let's now look at the kinds of thing that can provide convincing explanations of our surprising failures to do as our personality traits imply that we should.

III

To remind you of what I said earlier, I accept that all intentional action can be explained by appeal to beliefs and desires of the individual. However, belief-desire explanations, thin as they are, can be thickened out by appeal to factors that influence the way someone's mind works on a particular occasion, which can help to explain why someone has motives-beliefs and desires, intentions and so on-which are not what we would expect from his personality, and which, in turn, can help to explain his acting contrary to how we expected, and in a way that initially surprises us. These factors are not themselves entirely within what has been called the 'space of reasons': they so to speak bridge the divide between the mental and the physical; or rather, they throw into question the very idea of there being such a divide. I will put them into four broad categories.

First, there are states such as being drunk, being under the influences of drugs, having a bad cold, and being deprived of sleep.[6] A shy and retiring man goes to a cocktail party. He bumps into a woman whom he hasn't seen for many years. Suddenly, he is struck by the thought that he has been in love with her all this time, and, quite out of character, he tells her as much. We explain this by saying that he had had two glasses of champagne on an empty stomach; it was because he was a little bit drunk that he did what he did. What we would expect him to do, given his personality, is to have characteristically shy and retiring thoughts-'I couldn't possibly force my attentions on her; I hardly even know her.' Perhaps on this occasion he did have these thoughts, but normally, usually, we would expect them to have prevailed over any amorous inclinations he might have felt; and yet on this occasion they didn't prevail. You might think that this kind of case exemplifies weakness of the will. But weakness of the will is something of a moral notion: it is not straightforwardly obvious that the champagne *distorted* his thinking, and led him to allow *inappropriate* deliberative influence to his amorous inclinations: thinking and acting out of character because of drink and having your thinking distorted because of drink are not necessarily the same thing. Maybe she will turn out really to be the love of his life, in which case, in retrospect, we might come to agree that, without the champagne, being such a shy and retiring person, he would have missed out on the chance of his life. The champagne helped him to relax, to properly lift up his eyes to her beauty and her talents, and to enable him to see her as she really is. As Iris Murdoch put it, 'By opening our eyes we do not necessarily see what confronts us. We are anxiety-ridden animals. Our minds are continually active, fabricating an anxious, usually self-preoccupied, often falsifying veil which partially conceals the world.'[7]

My second category is emotions like being angry and being jealous, which, although capable of being justified or grounded in reasons, can influence thinking in ways that lead people to act contrary to what we would otherwise expect and in an unjustified way. You shout at your child for not sitting up straight in his high chair, in spite of being a caring and loving parent. You do this because you are angry. Your shouting at her can also be given a belief-desire explanation: you wanted her to sit up straight, and you believed that shouting 'Sit up straight!' at her was the best means of getting her to do this. But why did you, a caring and loving parent, have these

thoughts, so unusual for someone like you? Because you were unjustifiably angry, and because shouting at people who won't do what you want is characteristic of angry behaviour.

My third category is moods, relatively short-term states like being depressed, tense, irritable, full of unlocated sexual desire. Her not going to work today, in spite of being a diligent and hardworking person, and much to the surprise of her bosses, who has expected her to be there, can be explained by appeal to the fact that she felt depressed when she woke up this morning. Her being in this mood-depressed-helps to explain why she had the thoughts that she did, and thus goes beyond the belief-desire explanation. Perhaps what went on in her mind at the time was the thought that she might not be able to handle this large and important deal, and the thought that if she didn't turn up to work, then someone else would take the deal off her hands. Our thicker explanation-she was depressed-points towards an explanation of why she had those thoughts in spite of being normally so diligent and hardworking. Moods, in turn, can explain emotions, and they often do so without justifying them: for example, you got

unjustifiably angry with your child because you were tense and irritable.

Sometimes, and this is my fourth category, the explanation of an action that surprises us appeals to the influences of the particular situation that the individual finds himself in: his being in that situation brings about a surprising influence on thinking. The literature in social psychology these days is replete with examples which bring into question the reliability of personality traits in certain situations. I will mention just one.

In the summer of 1971, Philip Zimbardo, a psychologist at Stanford University, carried out an experiment on the effect of imprisonment. The account that follows is either direct citation from Zimbardo or paraphrase. He and his colleagues took a sample of 24 college students from the U.S. and Canada who happened to be in the Stanford area and wanted to earn $15 a day by participating in a study. On all measured dimensions, these students were typically healthy, intelligent, middle-class males. The participants were arbitrarily divided into two groups by a flip of the coin. Half were randomly assigned to be guards, the other to be prisoners.

The 'prisoners' were brought to the 'jail' (in fact the converted basement of the psychology department) one at a time and greeted by the warden, who conveyed the seriousness of their offence and their new status as prisoners. Each prisoner was systematically searched, stripped naked, and then deloused and issued with a uniform, a heavy chain bolted on the right ankle and worn at all times, rubber sandals, and a stocking cap made from a woman's nylon stocking. The guards were given no specific training on how to be guards. Instead they were told that they were free, within limits, to do whatever they thought was necessary to maintain law and order in the prison and to command the respect of the prisoners. As with real prisoners, the prisoners in this experiment expected some harassment, to have their privacy and some of their other civil rights violated while they were in prison, and to get a minimally adequate diet-all part of their informed consent agreement when they volunteered.

Because the first day passed without incident, Zimbardo and his colleagues were surprised by, and totally unprepared for, the rebellion which broke out on the morning of the second day. The prisoners removed their stocking caps, ripped off their numbers, and barricaded themselves inside the cells by putting their beds against the door, taunting and cursing the guards. The guards met force with force: they used fire extinguishers to get the prisoners away from the doors; they broke into each cell, stripped the prisoners naked, took the beds out, forced the ringleaders of the prisoner rebellion into solitary confinement, and generally began to harass and intimidate the prisoners. Once the riot was under control, they stepped up the violence and intimidation techniques. Every aspect of the prisoners' behaviour fell under the total and arbitrary control of the guards. Even going to the toilet became a privilege. Prisoners were often forced to urinate or defecate in a bucket that was left in their cell, and sometimes the guards would not allow prisoners to empty these buckets, so that the prison began to smell of urine and faeces. After a rumoured prison break, which in fact turned out to be false, the guards again escalated very noticeably their level of harassment, increasing the humiliation they made the prisoners suffer, forcing them to do menial, repetitive work such as cleaning out toilet bowls with their bare hands.

After five days, prisoners were seen being marched on a toilet run, bags over their heads, legs chained together, hands on each other's shoulders. At this point, long before the scheduled end of the experiment, Zimbardo was made to realise the psychological damage that the experiment was causing to the participants, and he quickly brought it to a close. To quote Zimbardo, 'We had created an overwhelmingly powerful situation-a situation in which prisoners were withdrawing and behaving in pathological ways, and in which some of the guards were behaving sadistically. Even the "good" guards felt helpless to intervene, and none of the guards quit while the study was in progress'; '...not once did any of the so-called good guards ever contest an order by a sadistic guard, intervene to stop or prevent despicable behaviour by another guard, or come to work late or leave early'. He reports that, amongst both guards and prisoners, '[N]one of our preliminary personality tests were able to predict this behavior. The only link between personality and prison behavior was a finding that prisoners with a high degree of authoritarianism endured our authoritarian prison environment longer than did other prisoners.'[8]

Given this dissociation of behaviour and personality, it would be a mistake to make an inference from a given type of behaviour on this single occasion, such as the brutal behaviour of many of the guards, to a brutal personality trait, and then appeal to that trait to predict and retrodict further brutal behaviour across a range of different kinds of situation. It isn't at all clear how to explain what these people did in this prison experiment, but I have no doubt that a simple 'brutal is who brutal does' is far too quick. The many experiments in social psychology show both how mistaken this is, and how prone we are to make the mistake.

Being drunk, being angry, being depressed, becoming a prison guard in particular circumstances: a large part of our everyday psychology is concerned with thicker explanations that appeal to factors such as these. And factors from each of my four rough categories can interweave in a complex network of narrative explanation: emotion, such as anger, can be explained by mood, such as being tense and irritable, as we saw earlier; being tense and irritable can be explained by lack of sleep and by situation, such as being harassed by a domineering boss; and so on. These narrative explanations explain through showing, sideways on so to speak, why people's patterns of thinking-their occurrent thoughts and choices-are as they are, and why they come to do things which surprise us. We find out, through our own experience, through literature and the arts, through reading newspapers, what characteristic influences various kinds of factors have. Being drunk often explains why people give way to temptation; we all know it can be a terribly expensive thing to go shopping after a boozy lunch. Being angry often explains why people do things like shout at their child or throw the wrong flavour of ice-cream on the floor; we all know the best thing to do when you're angry is to count backwards from a hundred. Being depressed often explains why we shirk from getting down to our work; we all know this too. Being thrown into an unusual situation without proper preparation, such as becoming a prison guard, can often explain why you behave as you shouldn't. It's all very human and understandable.

IV

What practical conclusions can be draw from all this? I would like to mention two, both of which bear specifically on character, rather than more widely on personality. The first concerns judgement. Let's say that some influencing factor leads you to act out of character in ways that were unexpected, and that this involved acting badly. We all thought, as did you, that you had some virtue-courage, say-and yet when you come to be tested you failed to act as you should have done. At the crucial moment, you acted to save your own life, in disregard of the others whom you should have helped.

Now, what I've been saying suggests two things. First, and this I mentioned earlier in relation to the Zimbardo prison experiment, we shouldn't simply brand you as having a bad character, as being a coward, 'cowardly is as cowardly does', just because of this one action. It's quite possible that you really are a courageous person, but less than fully virtuous, so courageous action isn't guaranteed on every occasion. After all, only flat characters are fully virtuous, and you are a round character. Secondly, we are round characters too, so we cannot rest complacently in the secure knowledge that our own virtue will smoothly generate virtuous action in all circumstances. So when we hear of what you did-that so-called cowardly action-we ought to reflect on the fragility of character, and acknowledge that we too might have done what you did if we had been in your circumstances. For who *really* knows what we might have done? As Marlow, Joseph Conrad's narrator in *Lord Jim,* said, 'Let no soul know, for the truth can be wrung out of us only

by some cruel, little, awful, catastrophe'.[9]

However, it is consistent with our thinking that we might have done the same as you in those circumstances, or even that just about anyone would have done the same as you in those circumstances, also for us to think that what you did was wrong, that you are responsible for what you did, and that you are to blame for it. Similarly, you should blame yourself for what you did. You should not hide behind the influencing factor-being a participant in the prison experiment for example-as an excuse to get you off the hook; as J. L. Austin said, some excuses get us on the hook rather than off it.[10]

But-and this is a point I cannot argue for here-your blame of yourself should not be misdirected, in remorsefully-and remorselessly-*looking backwards,* back to that one moment in your life, that one moment of failure of character. Without dodging responsibility, and without avoiding blaming yourself, you can focus your thoughts not, remorsefully, on the past, and on what you did wrong at that single, fateful moment, but on the future, and on changing yourself so that you won't make the same mistake again. As Nietzsche so brilliantly put it: 'Never yield to remorse, but at once tell yourself: remorse would simply mean adding to the first act of stupidity a second. - If we have done harm we should give thought to how we can do good'.[11]

The second practical conclusion, also bearing on the fragility of character, is that we should be circumspect about the reliability of our character and our motives on any particular occasion. The real springs of human action are a mystery, as much our own actions as those of others; and sometimes more so. Because we have names for things-'motive', 'deciding', 'willing', and so on-we too readily conclude that there is something very clear and precise that the names stand for. We may be able to deliberate about our motives, decide what to do, and later explain or make sense of what we have done-using names for motives, for deciding, for willing. But still, to quote Nietzsche again:

> ... at the moment when we finally do act, our action is often enough determined by a different species of motives than the species here under discussion, those involved in our 'picture of the consequences'. What here comes into play is the way we habitually expend our energy; or some slight instigation from a person whom we fear or honour or love; or our indolence, which prefers to do what lies closest to hand; or an excitation of our imagination brought about at the decisive moment by some immediate, very trivial event; quite incalculable physical influences come into play; caprice and waywardness come into play; some emotion or other happens quite by chance to leap forth; in short, there come into play motives in part unknown to us, in part known very ill, which we can *never* take account of *beforehand*.[12]

One might almost think that Nietzsche had in mind the recent experiments in social psychology when he wrote that. The idea, then, is this: if we can't be sure of our own motives on an occasion, or of our character, then we would do well to realise this in advance of the moment of action, and to plan accordingly. This is what I mean by circumspection-as the OED has it, 'attention to circumstances that may affect an action or decision; caution, care, heedfulness'. Like Odysseus, who had his crew tie him to the mast so he couldn't sail towards the tempting Sirens, we should sometimes plot against our future selves, by putting things in the way to prevent our being unduly influenced to act out of character.[13] Circumspection in forward planning is, like strength of will, a kind of executive virtue. But it comes temporally prior to strength of will; without the proper circumspection about our motives and character, strength of will can both be not enough and come too late. Don't rely on your character or your strength of will to see you through when you are tempted; better to keep yourself out of the way of temptation, or to keep temptation out of your way.

There's an interesting implication of this. If you acknowledge, as I say you should, that you are less than fully virtuous, not wholly and completely reliable in your actions, even (in that telling expression) with the best will in the world, then, somewhat ironically, it is a mistake to ask yourself, in thinking about what is the right thing to do, 'What would the virtuous person do here?'. For the *really* virtuous person, if there were such a flat character outside the pages of a novel, wouldn't be led astray, and could therefore *blithely* go into these situations, in which the rest of us would be tempted. So we shouldn't do what the fully virtuous person would do. The better question to ask is, 'What would the virtuous person *advise* me to do here?'. The virtuous person may well wisely advise you to be circumspect about your character and motives, and, like Odysseus, to take steps in advance to make it harder for you to act out of character in ways that you may well later come to regret.[14]

We are all round characters, capable of surprising in a convincing way. Accordingly, we must not, like Conrad's Marlow, hope for the impossible, 'for the laying of what is the most obstinate ghost of man's creation, [for the laying] of the uneasy doubt uprising like a mist, secret and gnawing like a worm, and more chilling than the certitude of death-the doubt of the sovereign power enthroned in a fixed standard of conduct'.[15]

King's College London

1 *Intention,* Oxford: Blackwell, 1958.

2 See especially Donald Davidson's 'Actions, Reasons, and Causes', together with the other papers in his *Essays on Actions and Events,* Oxford: Clarendon Press, 1980.

3 *Aspects of the Novel,* London: Pelican Books, 1962, p. 73.

4 *Aspects of the Novel*, p. 81.

5 *The Concept of Mind,* London: Penguin, 1990, p. 117; first published by Hutchinson, 1949.

6 These are called 'non-rational influences on thinking' by Jane Heal in 'Replication and functionalism,' in M. Davies and T. Stone, eds., *Folk Psychology: The Theory of Mind Debate,* Oxford, Blackwell, 1995, pp. 45-59. I discuss these issues in *The Emotions: A Philosophical Exploration*, Oxford: Clarendon Press, 2000, pp. 167-75.

7 Iris Murdoch, *The Sovereignty of Good*, London: Routledge, 1970, p. 84.

8 Citations from http://www.prisonexp.org/, and from 'Reflections on the Stanford Prison Experiment: Genesis, Transformations, Consequences', Zimbardo, P. G., Maslach, C., & Haney, C., in T. Blass, ed., *Obedience to authority: Current perspectives on the Milgram paradigm,* Mahwah, NJ: Lawrence Erlbaum, 2000, pp. 193-237. See the website for more references.

9 *Lord Jim, A Tale*, Oxford World Classics, Oxford: Oxford University Press, 2002, p. 236.

10 J. L. Austin, 'A plea for excuses,' reprinted in his *Philosophical Papers,* 3rd edition, Oxford: Oxford University Press, 1979, pp. 175-204.

11 In his *Human All Too Human,* 'The Wanderer and his Shadow', Section 323, translated by R. J. Hollingdale, Cambridge: Cambridge University Press, 1986.

12 *Daybreak,* translated by R. J. Hollingdale, Cambridge: Cambridge University Press, 1982, Section 129.

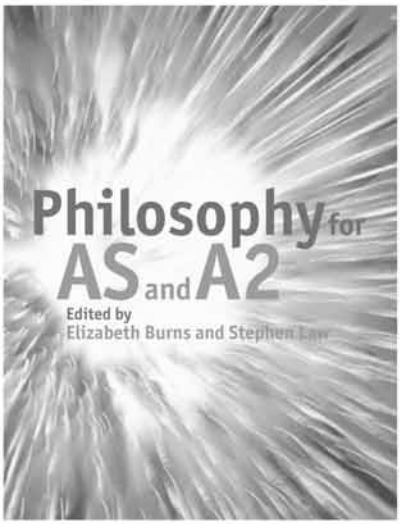
13 There is an interesting book about all this: *Ulysses and the Sirens: Studies in Rationality and Irrationality,* by Jon Elster, Cambridge: Cambridge University Press, 1979.

14 For discussion of the 'advice' and the 'example' or 'emulation' models, see Michael Smith, *Philosophy and Phenomenological Research* 55, 1995, pp. 109-31. The question people sometimes ask is 'What would Jesus do?', and it's interesting in this context to note that Jesus was tempted.

15 *Lord Jim*, p. 37.

New from Routledge Philosophy...

Special 20% discount for readers of the Richmond Journal of Philosophy!
Just quote the following code when placing your order: RICHMOND04

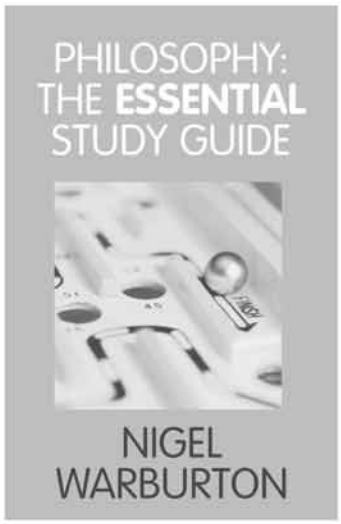


Philosophy for AS and A2

Edited by **Stephen Law** and **Elizabeth Burns**, Heythrop College, London

Philosophy is becoming ever more popular as both an AS and A2 level subject. This comprehensive survey of major philosophical themes in the syllabus fulfils a real need, and will be an invaluable tool for study.
– Professor Roger Trigg, Founding Chair of the British Philosophical Association

June 2004
Hb: 0-415-33562-0: £50.00
Pb: 0-415-33563-9: £14.99



Philosophy: The Essential Study Guide

Nigel Warburton, Open University, UK

Philosophy: The Essential Study Guide is a compact and straightforward guide to the skills needed to study philosophy, written by Nigel Warburton, best-selling author of *Philosophy: The Basics*. *Philosophy: The Essential Study Guide* is an indispensable guide for anyone getting to grips with their first philosophy course.

July 2004
Hb: 0-415-34179-5: £45.00
Pb: 0-415-34180-9: £7.99

Plus! Six new titles in the popular Routledge series *Thinking in Action*...

On Personality

Peter Goldie, Kings College, UK

The style is delightful, the non-philosophical works called on eclectic and quite fascinating ... the humour engaging ... the chapters beautifully balanced. *– Rosalind Hursthouse*

June 2004
Hb: 0-415-30513-6: £45.00
Pb: 0-415-30514-4: £8.99

On Cloning

John Harris, University of Manchester, UK

This book will change the way society views cloning. *– Julian Savulescu, University of Oxford*

June 2004
Hb: 0-415-31699-5: £45.00
Pb: 0-415-31700-2: £8.99

On Anxiety

Renata Salecl, University of Ljubljana, Slovenia and London School of Economics, UK

A deft application of psychoanalytic insights to very everyday and quite familiar human situations. *– Juliet Flower MacCannell, University of California*

June 2004
Hb: 0-415-31275-2: £45.00
Pb: 0-415-31276-0: £8.99

On Being Authentic

Charles Guignon, University of South Florida, USA

A very important book. We urgently need to be able to think clearly about authenticity, and Charles Guignon is uniquely prepared to say something important on this matter.
– Charles Taylor, author of The Ethics of Authenticity

June 2004
Hb: 0-415-26122-8: £45.00
Pb: 0-415-26123-6: £8.99

On Evil

Adam Morton, University of Oklahoma, UK

The author actually does what too few academics do: exhibit thought in action, not substituting the stale cite for an original idea. *– Fred Alford, author of What Evil Means to Us*

June 2004
Hb: 0-415-30518-7: £45.00
Pb: 0-415-30519-5: £8.99

On Humanism

Richard Norman, University of Kent, UK

A lucid account of humanism, which combines the virtues of a fairly balanced discussion and a passionate polemic. It deserves to become humanism s unofficial manifesto. *– Julian Baggini, Editor of The Philosophers Magazine*

June 2004
Hb: 0-415-30522-5: £45.00
Pb: 0-415-30523-3: £8.99

For direct credit card orders please call +44 (0) 1264 343071, e-mail **book.orders@routledge.co.uk** or visit **www.routledge.com**

For more information please call Anna Ridgewell on +44 (0) 20 7842 2128

Routledge
Taylor & Francis Group

Garrath Williams

[Two approaches to Moral responsibility]

Part 2

Introduction

In the first part of this article [Ed.- Which appeared in RJP Issue 6], I discussed two important understandings of moral responsibility. One approach links responsibility to the idea of free will, an idea which has, as we saw, Christian origins. On the other hand, Aristotle's account, predating Christianity by several centuries, views moral responsibility in the context of a moral community seeking to encourage and reinforce shared standards, where people make judgments of one another's character in the context of their on-going relationships with one another.

Since belief in heaven and hell is now rarely thought of, even by many who consider themselves Christian, one might suppose that our sympathies would naturally go to Aristotle's account. Yet this is not the case. The idea of moral responsibility so well captured in Kant's philosophy rests on a powerful notion of moral worth that continues to have strong intuitive appeal. Here I want to set out Kant's account in more detail, before suggesting that we can safely give up some of the intuitions that support his account of moral responsibility.

The Kantian account: freedom, intentions and control

The reason why so many people - even before they come to philosophy - feel the pull of the free will debate lies in the idea of moral worth we often associate with responsibility attributions like blame. Galen Strawson expresses the core idea as follows: 'if we have [true responsibility], then it makes sense, at least, to suppose that it might be just to punish some with eternal torment in hell, and reward others with eternal bliss in heaven' (1991: viii). Any such 'ultimate' merit or demerit is clearly a matter of strictly individual desert. If it were merely a matter of chance who went to heaven or hell - or who *would* do so, if those fates really existed - this would plainly be a matter of mere fortune. Such intense good or bad luck would make the world even more morally arbitrary than it already is. If such merit is to be fairly allocated, therefore, it needs to be seen as something that lies within individuals' own control. This line of thought, in turn, relies on what John Skorupski calls an 'ideal of pure egalitarian desert' (1999: 156). The idea is that we all equally possess such control, so that it makes sense to imagine everybody reaping an equally fair return on how well we exercise that control.

The greatest representative of this line of thought is Kant. For Kant, our moral worth - the goodness of our will - is gauged by how sincerely and persistently we have sought to do our duty. To do our duty may be much harder for some people, for instance, those who have violent passions or who were brought up with bad habits. But moral worth is not about results; it is about *the will*. We all have such a will, an ability to choose well, despite the fact that some of us face stronger counter-inclinations or more difficult circumstances. To truly judge a person's moral worth involves seeing past all the obstacles that their will has faced. Kant argues that this makes moral worth impossible for us to judge with any assurance; only God can see beyond all those things. Kant's main concern, though, is how we judge *ourselves.* To Kant it's no problem that we're never sure about others' wills, and the obstacles or benefits they have faced. The point is that we can never be sure of our own motivations: we must keep trying to do the right thing, and to do it *because* it is the right thing.

But what is 'the right thing'? What about those who are morally ignorant? Kant denies that there is such a thing as moral ignorance, in that we are all equally well able to see what we should do - if only we try. For Kant 'even the most hardened scoundrel' would act morally, were it not for the

opposing incentives of his inclinations and desires (*Groundwork*, 4:454). Kant has to claim this because otherwise he would not be able to justify condemning people who suppose they are doing the right thing, when in fact their acts are quite wicked. Adolf Eichmann, who we mentioned in the first part of this article, thought he could justify his actions (ironically, with a version of Kant's moral philosophy!). Yet no one doubts that he deserved the gravest condemnation for his crimes, even though many believe he was sincere in thinking his acts were defensible. Such examples show how implausible is the claim to equal moral knowledge. Even outside of such extreme cases, people's sensitivity to different moral considerations is highly variable, and is clearly shaped by up-bringing and environment. From an Aristotelian perspective, though, the realities of moral ignorance and moral disagreement pose no theoretical problems. In fact, they provide an important justification for praise and blame in terms of mutual accountability. Responsibility attributions are important to moral learning, by communicating how we have met or failed to meet moral standards. Because the Kantian account goes inward, to *my* scrutiny of *my* motives and intentions, it ignores this crucial *educative* aspect of responsibility attributions.

This striving to improve my intentions is tightly bound up with the idea of desert. It is a basic and very appealing intuition of Kant's ethics that *happiness should correspond to moral worth* - even though in the world we know this is invariably not the case (at least so far as we can judge people's moral worth at all). For this reason Kant claims we are rationally and morally committed to faith in a *future* world, where virtue will be rewarded. (Kant is completely silent on whether the wicked should also be punished; again, this relates to his principal concern, the striving of the self rather than the judgment of others.) Few modern Kantians follow Kant in thinking we must have faith in the compensatory schema of an after-life. But many people remain deeply attached to the idea of a distinctively *moral* sort of worth that Kant so consistently articulated.

As we have seen, such worth is measured by the sincere intentions that are within *everyone's* control. This conception corresponds quite plausibly to several features of moral responsibility. We judge the intent behind people's actions, rather than the often haphazard results of our actions. We take account of people's circumstances, and judge less harshly where these place hard or immoral pressures on people. We also, quite often, feel that allowances should be made for the effects on character of abusive or deprived upbringings. In each case, we can interpret these concessions in Kantian terms - as drawing a distinction between the person's will and the obstacles of circumstance, thus keeping our moral evaluation to what is within a person's control. There are, however, good reasons to doubt whether this Kantian interpretation is really the best account of these intuitions.

First, it is a commonplace that we expect people to take responsibility for things they didn't intend. This is not only in those cases where we judge that someone should have formed their intentions more carefully. Certainly we judge the negligent driver who causes an accident more harshly than a driver who was careful but nevertheless caused an accident. But even in the latter case, we expect the driver to bear important responsibilities. (The problem that many of the things which attract moral merit and demerit are wholly or partly outside of individual control is known to philosophers as the problem of moral luck. Kantians *deny* that our moral worth could ever depend on luck, whether this is a matter of our upbringing or the circumstances we encounter.)

Second, when we are allocating things on the basis of desert, this is usually not to do with what the individual concerned has been able to control. The usual criterion for a job appointment is merit, measured in terms of the demands of the post: on this basis, the best applicant deserves to be offered the job. The merits concerned may be matters outside anybody's control, for instance, the looks of a fashion model, or they may be only partly related to those qualities of character that we assume people have some control over. Why, then, should control be so important in allocating '*moral* desert'? (Perhaps it will be said that it is a defining feature of moral desert that it relates only to what is within our control. Nonetheless, the question remains: why should we think this to be so?)

Not least, some of the concessions to circumstances that the Kantian story permits can be accounted for in quite different terms. We might take account of *intention*, for example, because an intended action reveals character especially clearly. But if we always look back to the will, we face grave difficulties in accounting for the moral evaluation of unintended actions. Just the fact that a person did *not* intend to do something may say something important about her character. Negligence is a case in point: the person lacks appropriate habits of attention. Inconsiderateness is another. The person does not *intend* to be selfish; rather, he *fails* to form intentions with regard to the effects his conduct has on others. In other words, the Kantian account, being focussed upon control, stumbles when a person fails to control herself. An account based on character, on the other hand, cares both about a person's intentions *and* the things she does out of mere habit or inattention.

These are reasons for caution but not definitive objections to the Kantian notion of moral worth. One might, for instance, claim that responsibility in practice represents a combination of Kantian and *consequentialist* tendencies. We want to be fair to people's moral worth, but we also need to take account of the consequences of how we allocate responsibility. Thus we usually make concessions to intentions, circumstances, up-bringing in the name of individuals' moral worth. But sometimes we have to be more concerned with overall results than individual fairness - for instance in public systems of punishment and liability. Here, there might be good reasons to punish or reward regardless of inner striving, especially as it is so difficult to know about a person's will.

The idea of worth

The notion of moral worth central to Kant's account is probably what one writer on ancient Greek ethics - AWH Adkins - had in mind when he said, 'We are all Kantians now.' (1960: 2) Kant's idea attractively reconciles two broad value judgments: (i) the moral egalitarian idea that all persons are moral equals by virtue of having freedom to choose morally; and (ii) the idea that responsibility relates to desert, so that people can nonetheless be judged very differently - some being blamed for their lives and characters, others praised. We have seen that it is not really plausible to think that people have an equal ability to choose well. But the deepest difficulties for a Kantian account turn on the idea of a moral worth, where moral responsibility is understood in terms of a person's 'ultimate' merit.

To begin with, contrast Kant with Aristotle. Aristotle makes no claims about a person's ultimate merit or demerit. People might be vicious or virtuous in various ways, and there might be rare paragons who possess a comprehensive set of virtues (yes, these are philosophers!). Naturally we would not want to associate with the vicious, and naturally we will want to condemn their vices in no uncertain terms. It might help them to learn to do better, and it may caution others against them, and it should reinforce our own and other people's sense of what character traits are desirable. But for Aristotle there is no sense that the vicious are earning a lasting form of discredit that should condemn them in the eyes of an ultimate judge. If the vicious person were to protest to Aristotle that the condemnations he faced were unfair, perhaps because his character had been formed by his vicious parents, one suspects Aristotle would be rather unmoved. Life isn't fair, he might say, and we certainly won't make it fairer by pretending some vices are less real because of

their origin in early childhood, let alone because of their fixity within an individual's character. It may be unpleasant (he might continue) for you to hear this blame and condemnation - indeed, I'm glad that it is, because at least it shows that you are not so vicious that you don't care about others' opinions of you - but there are other matters at stake here, above all the standards and expectations which regulate all our lives together.

This line of thought will always be unsatisfying if we think that our praise and blame of one another reflect peculiarly deep truths about people's moral deserts - so that there are some people who are *evil* and deserve to go to hell, others who are good and deserve to go to heaven. But this way of thinking poses a series of deep problems.

In the first place, it is very difficult to know what to make of it within a secular framework, where religious faith has become a personal matter. In particular, why should someone with no religious faith continue to believe in this idea of personal moral worth? We should not forget that there are several morally *disreputable* motives that can make this idea attractive. Revenge against wrong-doers, or hatred of them, might be justified by thinking that they really deserve these reactions. And Nietzsche made a famously cynical point: if we think of ourselves as free, and also see ourselves as leading good or at least blameless lives, then we can smugly take credit for our moral superiority over others. (Nietzsche was also rightly indignant at some early Christian writers, who claimed that one of the pleasures of heaven would be looking down on hell and the tortures of the damned.)

Second, and perhaps more important, such thinking is incoherent on its own terms. Ultimate deserts are, precisely, ultimate: they are for God to judge. Perhaps a priesthood might be thought of as having been granted some provisional right to judge, but such an idea has become increasingly unconvincing even to the devout. Again, an individual concerned with personal salvation will naturally wish to scrutinise himself to become worthy of God's esteem. But it seems over-ambitious, to say the least, to construe our praise and blame of one another as tentative or fallible versions of what God will say. (Perhaps the non-believer could say here: given that God does not exist, it is *our* task to ensure that people get their moral deserts. To this one can object: (i) As before, it is not clear why we should think that this 'ultimate' moral desert exists; (ii) Even if it does, why should rewarding and punishing it be of particular concern to us?) Although Kant's writings are littered with moral evaluations of people and their actions, his theoretical position is that we are in no position to judge the worth or deserts of others.

This leads to a third large problem in thinking in this way: it simply does not fit with what usually goes on when we hold one another accountable. In our resentment or indignation, we might want to lend our judgments as much force as possible, and so rely on any ideas that can lend our mere words greater authority ('You deserve to burn in hell!'). But when we're not over-egging the moral pudding, it is plain that praising and blaming are fairly mundane practices. Holding each other responsible assumes that we all get things wrong in one way or another, and that for various reasons some people get things badly wrong, and that we must - for just these reasons - seek or reinforce a shared set of standards and expectations. Moreover, while we do sometimes impute good or ill will to people, it is a childish morality that is preoccupied with who is 'good' and who is 'bad': our mature moral judgments are much more complex, multivalent and fine-grained than this.

Why does the idea of moral desert continue to preoccupy us?

Why, then, does the 'moral bank account' Feinberg ironically referred to, the 'true responsibility' Galen Strawson so memorably describes, continue to hold such sway over our moral sensibilities?

As we have seen, there is a *distributive* aspect to our responsibility attributions. If something has gone wrong, and several people have had a hand in the matter, we often go to great lengths to assess who bears what degree of responsibility - who should make recompense, who should apologise, or even who should be punished. At least in part, the question is certainly: who *deserves* to be held accountable? But does this necessarily point us toward the idea of lasting moral merit or demerit? The idea of a stain on our character, when we fail to acknowledge our guilt, or do something that cannot be made good, might support this way of thinking. In another way, the idea of 'recompense' suggests the image of an account, albeit one we can again make good. But for both cases a more mundane explanation is possible: that we judge something about the character of the person we are dealing with, from how they act and how they respond to others' responsibility attributions. How well does he understand the needs and interests of those around him? What sort of cooperative relations does it make sense to pursue with him? What can we trust him with? Not least, is this a person who will *take and accept responsibility?* (Note, further, that these things do not form a neat package: some people might do better on one count than they do on another - hence the many different virtues we ascribe to different people. This point might reinforce doubts about the single, 'ultimate' evaluation of a person's will as good or bad that is so important to Kant's account.)

Naturally, to judge such questions we need to appreciate what pressures were placed upon a person when he acted, how he understood the situation he was responding to, and special factors affecting his ability to deliberate and choose. Hence Aristotle's concern with force of circumstances, factual ignorance, and intoxication, and our more modern concern with mental illness. But on an Aristotelian account, the point is that *these factors alter the extent to which actions reveal the character of the person.* That they undermine the person's 'control' is true, *but subsidiary.* We can see this by considering how certain forms of bad character constitute a lack of control over one's actions - thus the person who is weak-willed or indecisive, for example. Here weak-willed, indecisive action reveals the person, in a way that action based on ignorance of the facts does not. This also makes good sense of Aristotle's, and our, ambivalence about coerced action. Plainly someone else, the coercer, can fairly be held responsible for the coerced person's deed; and if we were concerned solely with *distributing* responsibility that might be all we needed to say. But clearly our judgments do not come to rest so quickly: we also, habitually, evaluate the coerced person's conduct. Should she have done such and such in response to such a threat? Was giving way cowardly or prudent, feeble or tough-minded?

In other words, distributing one particular quality, blame or 'blameworthiness,' is not our only or even our real concern in cases of harm and wrong. In the first place, there is the judgment of character, which is *not* an exercise in distributive fairness. (When we are judging people in terms of their fitness and abilities to interact with others, our concern is extremely partial: above all, to evaluate those people who *we* might continue to interact with.) Second, the most important thing to be distributed is not so much *blame* but the resulting *responsibilities:* who should pay compensation, apologise, or - in the case of those who have manifested the most harmful sorts of irresponsibility - be punished. Certainly this exercise is guided by the distributive ideal of fairness, but only in part. Practicality, limited knowledge, and many consequentialist considerations also play important roles. Third, as Bernard Williams has

stressed, an important part of what we seek in praising and blaming is 'to recruit people into a deliberative community that shares ethical reasons' (1995a: 16). Thus when we blame we demand that someone take *more* seriously reasons to think or act one way rather than another - more simply, that he takes *our* reasons to be *his* reasons too.

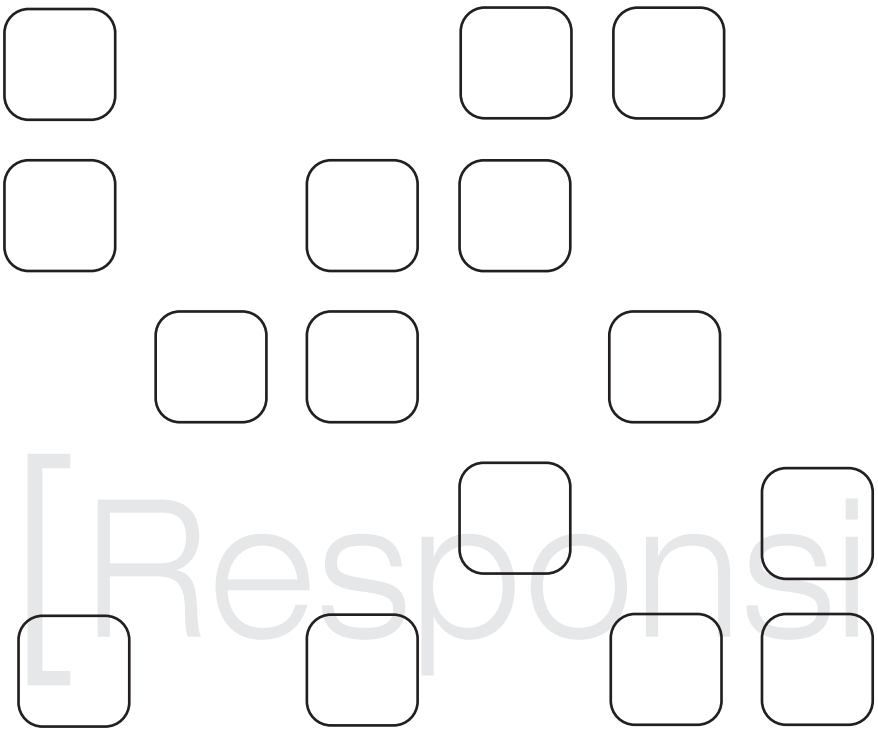
Simplifying both positions, we can say that what essentially separates Aristotelians from Kantians here is whether there is something 'more' at stake in our attributions of blame. Once character has been judged, responsibilities distributed and others encouraged to take on board our reasons and evaluations, Aristotle's story is at an end. From such a perspective the question is not, 'How can people be truly responsible for their deeds?' but rather, 'Why should a notion of "true responsibility" continue to preoccupy us?'

The Kantian will feel that something has been left out. The person whose character is judged wanting, the person who should make recompense, the person who failed to appreciate a reason to act differently - mustn't they have *earned* the blame if they are truly to *deserve* it? And how could we earn condemnation except by doings for which we are *truly* responsible? My own view, it will be clear, is that we should resist the temptations of such an 'ultimate' responsibility. We do not need such a notion to make sense of how we hold one another responsible - at least, not of how we hold people responsible when we are not being self-righteous or violently indignant. While one strand of Christianity may still tempt us to think of the righteous and of the damned, another embodies a rather deeper wisdom. Even as a person holds the wrong-doer responsible, it is possible to say, 'There but for the grace of God go I.'

Conclusion

Ideas about responsibility are usually presented in terms of a contest between two positions, compatibilism and incompatibilism. Incompatibilists accept the dilemma of free will versus determinism: responsibility depends on me controlling my actions, rather than other causal influences that operate upon me. Praise, but especially blame, make no sense if determinism is true. Compatibilists, on the other hand, want to insist that the causal well-orderedness of the universe is, precisely, compatible with our responsibility for our actions. But for most philosophers the question is *not* whether responsibility and causal well-orderedness are compatible, but *how*. In other words, to adapt Adkins's adage, 'we are all compatibilists now.'

However, the essential issue for any compatibilist position lies in the conception of responsibility it relies on - and this issue has been much less well-explored by philosophers than the metaphysics of freedom and determinism. I have contrasted two broad schools of thought that reflect large aspects of how we put responsibility into practice. When Adkins claimed that 'we are all Kantians now,' he was not referring to Kant's (*in*compatibilist) metaphysics but rather to our tendency to feel that responsibility attributions must have depth, that they reflect something about a person's 'real' deserts. Yet this position leads us to claims about control over the self, to the idea of choices that were 'really' our own and not the result of any external influence - a position that always threatens to bring us back to metaphysical freedom and to incompatibilism.



The roughly Aristotelian alternative sketched here owes much to Bernard Williams and his critique of a distinctively modern notion of 'morality,' a notion most systematically expressed in Kant's philosophy but also expressed in alternatives such as utilitarianism. Williams argues that these ideas neither make sense on their own terms, nor do they make sense of what we actually do when we do engage in attributions of responsibility. As we have seen, Aristotle's account of praise and blame is based on: (i) how far acts reveal character; (ii) the fair distribution of responsibilities to act; and (iii) the attempt to exchange reasons, share standards, and maintain relationships with those whom we judge - and who judge us in turn. This account involves no pressure to think of people as responsible for their acts in any 'deep' sense - the sort of deep sense that would make eternal punishment or eternal reward intelligible. The basic facts of compatibilism's causally well-ordered world are that we can rely on no-one to judge such deserts - except ourselves - and that we can rely on no-one to mete out such rewards or punishment - except ourselves. In this situation we must do our best to relate to each other as best we can, and do our best with one another when we do not relate to others as well as we might. Fortunately this does not depend on an untenable idea of 'true' responsibility - only that we encourage people to take responsibility for their actions, and be prepared to do the same ourselves.

Acknowledgements

I would like to thank David Archard for comments on a version of these articles. Much of this material also appears, in rather different form, in an entry on 'Praise and blame' for the Internet Encyclopedia of Philosophy: www.utm.edu/research/iep.

References and further reading

Adkins, AWH (1960) *Merit and Responsibility,* Clarendon Press, Oxford

Aristotle *Nicomachean Ethics* (the most readable translation is Roger Crisp's, Cambridge University Press, Cambridge, 2000)

Feinberg, Joel (1970) *Doing and Deserving: Essays in the Theory of Responsibility* (Princeton University Press, Princeton NJ) - a set of classic essays on responsibility for action

Fingarette, Herbert (1967) *On Responsibility* (Basic Books, New York) - another set of classic essays, including the argument that blame is intelligible only insofar as it addresses a person's pre-existing concern for others

Kant, Immanuel (1784) *Groundwork to the Metaphysics of Morals* (the best translation is Mary Gregor's, Cambridge University Press, Cambridge, 1998)

Korsgaard, Christine (1996) 'Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations' in her *Creating the Kingdom of Ends* (Cambridge University Press, Cambridge) - a sophisticated Kantian account of holding people responsible

Midgley, Mary (1991) *Can't we Make Moral Judgments?* (Bristol Press, Bristol) - an accessible argument for the importance and inescapability of moral judgment

Skorupski, John (1999) 'The Definition of Morality' in his *Ethical Explorations* (Oxford University Press, Oxford)

Smiley, Marion (1992) *Moral Responsibility and the Boundaries of Community: Power and Accountability from a Pragmatic Point of View* (University of Chicago Press, Chicago) - criticises conventional discussions of freedom and determinism, claiming that they fail to investigate the idea of responsibility

Strawson, Galen (1991) *Freedom and Belief* (Clarendon, Oxford)

Strawson, Peter (1974) *'Freedom and Resentment' in his Freedom and Resentment and Other Essays* (Methuen, London) - this famous essay resituates the free will debate by highlighting the importance of 'reactive attitudes' such as resentment to interpersonal relations

Williams, Bernard (1993) *Shame and Necessity* (University of California Press, Berkeley CA) - a sustained argument that the ancient Greeks had a nuanced and sophisticated account of responsibility attributions

- (1995a) 'How Free does the Will Need to be?' in his *Making Sense of Humanity and Other Philosophical Papers, 1982-1993* (Cambridge University Press, Cambridge)

- (1995b) 'Voluntary Acts and Responsible Agents,' in his *Making Sense of Humanity*

Christopher Hamilton

[Sexual Desire]

Some Philosophical Reflections

Compared with many issues, philosophers have not written a great deal on the nature of sexual desire, and what they have written on it is has not always been very enlightening. One reason for this is that philosophical discussions of sexual desire have often been fairly heavily moralised. That is to say, philosophers and others have often presented a particular *moral* conception of sexual desire as if it were an understanding of sexual desire as such, thus distorting our understanding of this phenomenon (or series of phenomena) of human life. Certainly, it must be granted, I think, that it is doubtful that one can arrive at an account of sexual desire which is wholly free of moral concern, but this just means that if we are aiming to give an account of sexual desire that is honest and realistic we should aim to be very sensitive to the moral notions that we employ in doing so. In what follows, I discuss some key philosophical theories of sexual desire, in which theories moral concerns are present in differing ways, and then add some thoughts of my own.

In my view, the most profound philosophical account of sexual desire is that provided by Jean-Paul Sartre in *L'Être et le néant [Being and Nothingness].* Sartre begins his discussion of sexual desire by dismissing the view that sexual desire is a desire for pleasure. He does so since he claims that, if such desire were a desire for pleasure, then it would be impossible to make sense of how it is that such desire could come to 'attach' itself to an object, that is, to another human being. Crudely put, if desire were desire for pleasure, why would masturbation not be enough?[1] What, then, according to Sartre, does one want in experiencing sexual desire? We can approach his discussion by considering his reflections on the nature of the caress. Such a caress - it may be a caress of the hand or the eye - constitutes, says Sartre, an attempt to *incarnate* the other.

The other, he says, is born as flesh under my caress, whence the idea that I want him or her to be overwhelmed by his or her body: 'Desire is the attempt to strip the body of its movements as of its clothing and to make it exist as pure flesh'.[2] If the other responds to my caress then this person will experience his or her arousal as 'troubling', as 'clogging' consciousness. Yet, at the same time, my experiencing my own desire is felt by me in the same way, and I, too, in responding to the caress of the other, am born as flesh for him or her.

We can put Sartre's account in this way. If I desire *you*, I do not desire your flesh. Rather, I desire you in *your flesh.* It is you I want to exist *as* flesh for me. I want to possess you, not as mere flesh, but *through and as revealed in your flesh.* For Sartre, this 'you' is your freedom, for Sartre identifies the self and freedom. But one does not have to accept that identification to see the power of Sartre's account. We are embodied creatures, and our consciousness of that is crucial to our life. When we share a meal, or walk together, or talk together, we can only do so in the way we do because we are embodied. But if we share a meal with each other, we are not interested in one another *as* embodied.

However, if I desire you sexually, then I *am* interested in you *as* embodied. This is why being the object of sexual desire can be so compromising: suddenly to be aware that another desires one fills one with a consciousness of one's being an embodied creature. One is aware of one's flesh as revealing who one is, and as being the focus of the other's interest in one.

For Sartre, the fact that sexual desire has this kind of structure, i.e., that it is a desire for a person in his or her flesh, means that it is doomed to failure. Remember that, for Sartre, in desire I want to capture your freedom in your flesh. But if I manage to possess your freedom on the surface of your flesh, then you are, of course, no longer free. For if I possess your freedom, then I hold it captive, and, in holding it captive, it is clearly no longer free. Thus, if I achieve what I want in my sexual desire for you, namely, possessing your freedom, then I have thereby thwarted or frustrated my own desire. But you, too, are caught in the same process in your desire for me: if you capture me in my freedom, then I am no longer free, and you have failed to achieve what you want to achieve. This means that in our desire for each other, we are experiencing a conflict with ourselves and with each other. We can neither of us get what we want, and yet in our desire we struggle to do so. This is why Sartre claims that orgasm cannot be the aim of desire. Rather, orgasm signals the frustration of desire, since it is, so to speak, the point at which the failure to capture the other in his or her flesh becomes manifest.

Sartre's account, only the bare bones of which I have provided here, clearly captures something central about the nature of sexual desire. For even if we do not accept his identification of the self and freedom, there is, it seems to me, something right about the idea that sexual desire is doomed to a peculiar kind of frustration. Of course, all of our pleasures can fail to bring satisfaction: nothing is more common than to satisfy a desire and remain unsatisfied oneself. But the point about sexual desire goes deeper than this: sexual desire seems in a special or peculiar way doomed to frustration. But if that is not, as Sartre in his account proposes, because of the identification of self and freedom, why is it? Here is a suggestion. Sexual desire seems to be a deeply unstable desire. On the one hand, it is roving, largely undiscriminating about the individuals to whom it attaches itself, restless: one wants 'woman' or 'man'. On the other, it can be especially excited by, and become fixated upon, a specific individual. This lends sexual desire a strange fragility: for, in desiring a given individual, one also desires him or her as man or woman, as a representative of the male or the female sex. There accordingly seems to be a way in which what one wants in the sexual act is two things that one cannot have: one wants this individual man or woman and one wants all men or all women. That is, one wants *all* men or *all* women in and through this one individual. But this is impossible. And this is perhaps part of the explanation for the fact that sexual desire can be so imperious and desperate. It may also be the reason why one of the most recurrent sexual fantasies is that of not knowing who one's sexual partner is.

But Sartre's account seems weak in one crucial way. He starts, as we have seen, from the idea that sexual desire could not be desire for pleasure since if that were so then we could not explain how desire attaches itself to another. But this seems mistaken. The reason for this is that the pleasure that comes from sex with another might simply be more intense or more varied than the pleasure that comes from masturbation. We could thus explain how desire attaches itself to another by saying that it is this intense or multifaceted form of pleasure that is wanted in desire, and that this can only be satisfied by actually having sex with another, whence desire attaches itself to another.

Roger Scruton would disagree with the last point. He has argued that any instance of sexual desire possesses an individualising intentionality. By this he means that sexual desire is founded upon the thought of the other as the specific individual he or she is.[3] That is, there can be no sexual desire which exists and then 'attaches' itself to a specific individual. Hence, according to this account, if a man desires two women at the same time, he will be experiencing two different desires, each of which will be a desire for one of the two women. From this account it also follows that there cannot be any such sexual desire as an unfocused desire for no particular man or woman. Scruton considers the case

of the sailor storming ashore with the thought 'woman' in his mind: he might be thought to desire a woman, but no particular woman. Scruton claims that this is not so: until the sailor actually meets a specific woman he desires, he desired no woman; he was rather in the condition of desiring to desire.[4]

Such a view of sexual desire has to find an adequate response to such phenomena as that of Casanova, described by Stefan Zweig:

> His passion, flowing away at the purely erotic level, knows nothing of the ecstasy of uniqueness. We need have no anxiety, therefore, when he seems reduced to despair because Henriette or the beautiful Portuguese lady has left him. We know that he will not blow out his brains; nor are we surprised to find him, a day or two later, amusing himself in the first convenient brothel. If the nun C.C. is unable to come over from Murano, and the lay-sister M.M. arrives in her place, Casanova is speedily consoled. After all, one woman is as good as another![5]

Scruton writes: 'If John is frustrated in his pursuit of Mary, there is something inapposite in the advice "Take Elizabeth, she will do just as well."'[6] Not, apparently, if one is Casanova! It seems, then, that Scruton has two options. Either he could insist that he has provided a true account of sexual desire, in which case Zweig has totally misunderstood and misdescribed the case of someone like Casanova, and, indeed, that a lot of what looks like sexual desire where what is desired *is someone or other* is not really sexual desire after all since it does not display an individualising intentionality; or he could say that such cases display sexual desire all right, but in a perverted or otherwise morally unacceptable form. In fact, Scruton seems to waver between the two, for, although, as we have seen, he claims that in cases such as that of the sailor the man in question experiences no sexual desire until he comes into contact with the woman he desires, he also grants, at the end of his book, and looking over his argument as a whole, that 'my analysis has included a large prescriptive component'.[7] In other words, he seems to concede that his analysis is not an analysis of sexual desire as such but a moral view about the best form that sexual desire can take. It is, in other words, a moralised account of sexual desire.

I do not think, then, that Scruton's account is wholly plausible as it stands. However, it seems to me clear that what Scruton is trying to do is to give an account of sexual desire that does justice to the fact that there can be deeper and shallower expressions of such desire. Indeed, it seems to be the case that many people long for their sexual desire to be provided with deeper forms of expression. But some accounts of sexual desire do not seem to be able to make sense of this. One such is that provided by Igor Primoratz, who has argued that sexual desire 'is sufficiently defined as the desire for certain bodily pleasures, period'.[8] The reason that such an account of sexual desire makes it hard to see how such desire is capable of finding deeper forms of expression in human life is that it assimilates sexual desire to something like the desire to scratch an itch, and the possibilities of a deepened understanding of itch-scratching are severely limited, to say the least. This is not to say that only deepened forms of expression of sexual desire are morally legitimate, or anything like that: it is merely to say that any account of sexual desire must be able to make sense of the possibility of those deeper forms of expression.

In any case, Primoratz' account of sexual desire has some odd consequences. It leads, he argues, to the conclusion that any putative sexual act which is devoid of pleasure for the person engaged in that act is not, after all, a sexual act at all. Thus he claims that a prostitute who gains no pleasure from intercourse with a customer is not engaged in a sexual act (whereas the customer is). Further:

> As for the couple who have lost sexual interest in each other but still engage in routine coitus, the less pleasurable it gets, the less valuable it is as sex. If, at some point, it becomes utterly bereft of sexual pleasure, would it be so odd to say that they were performing acts that for most people ordinarily involve at least a modicum of sexual pleasure, but that *they* were merely going through the motions, that *for them* there was no sex in it any longer?[9]

One might suspect that Primoratz is not, after all, just trying to tell us what sex is, but prescribing a particular form of it, that is, one through which one experiences as much pleasure as possible. For he clearly believes that the less pleasurable sex is, the less valuable it is. Still, leaving that aside, it does, surely, seem odd to suppose that the bored couple in Primoratz' example are not actually engaged in a sexual act. One might as well say that what it is to feel hunger is to have a desire for certain bodily pleasures so that if one eats something utterly bland which fails to fill the stomach (modern mass-produced strawberries, for example) one is not really eating at all.

In fact, I do not think that Primoratz need deny on his account that the prostitute or the bored couple are engaged in sex even if they get no pleasure from such acts. His view expresses a confusion between sexual desire and sexual acts. One is, after all, still eating if there is no pleasure in doing so. The prostitute might not, indeed, possess any sexual desire for her clients, but it does not follow from that that she is not engaged in sexual acts with them. The same may be the case for the bored couple. In the same way, I might for some reason have no hunger, no desire for food, yet still be eating.

So far, then, we have seen that three key philosophical theories of sexual desire have weaknesses, though I certainly would not deny that they each capture some part of the truth about some individuals' experience of sexual desire. But if we were to try to find some fundamental reason why they are not complete as accounts of sexual desire, why they do not do enough to open up a deepened understanding of sexual desire, then I think that we would have to note that central here is that none of them makes anything of the connection between sexual desire and procreation. And we can see that this connection is crucial by the simple reflection that a species of creature which had all our experiences of sexual desires but in whom sexual desire had no connection with procreation would have a profoundly different understanding of sexual desire from the one we have. As so often in philosophy, the real problem is to find a helpful way of expressing this point.

At one point D. H. Lawrence writes:

> Sex is the balance of male and female in the universe, the attraction, the repulsion, the transit of neutrality, the new attraction, the new repulsion, always different, always new. The long neuter spell of

> Lent, when the blood is low, and the delight of the Easter kiss, the sexual revel of the spring, the passion of mid-summer, the slow recoil, revolt, and grief of autumn, greyness again, then the sharp stimulus of long winter nights. Sex goes through the rhythm of the year, in man and woman, ceaselessly changing: the rhythm of the sun in his relation to the earth.[10]

It goes without saying that many, if not most, do not share this view of sex, wonderful though it is. And there are lots of ways in which one might pursue or develop or respond to the thoughts Lawrence articulates. For our purposes what is important is that Lawrence connects sex to the natural cycle of life, and does so in such a way as to express a sense of the wonder and mystery of sex. But if we ask ourselves how it is possible to see sex in this way, then I think that we shall not be able long to resist the thought that it is the fact that sex is related to conception and procreation that allows us to do this. For it is *this* fact about it which most immediately and forcefully connects it to the notions of corruption and regeneration and hence allows it to be brought into contact with our sense of the natural cycle of the seasons. And if, as we do, we can wonder at that cycle, at its

utter familiarity together with the strangeness that each spring green shoots sprout from what looks like dead wood, we can also see why it is that we can wonder at sex, at the strangeness of a force at once so familiar and yet *unheimlich* - this incomparable German word, which means 'uncanny' or 'spooky' or 'frightening', captures the sense of something's not being like that which one meets with at home *[Heim]*, that which is unfamiliar or upsets one's ingrained and habitual ways of dealing with things.

We could perhaps get at the significance of procreation for an understanding of sexual desire in another way. Many people experience a sense of wonder and mystery at the birth of a child. And this very sense can cast in a certain light the sexual act which directly led to this birth, can remind us of the strangeness and mystery of sex. But to speak here of a reminder is not to suggest that anyone might actually have forgotten anything, for we are all familiar with the fact that sexual desire has its own demands and needs which well up and grip us in ways we cannot fully fathom, and that it attaches us to people in ways we cannot properly comprehend. We all know that sex, where what is craved is so clear and yet weirdly elusive, seems at once completely natural and an intrusion from another world into our daily activities.

The issue is rather that of such knowledge becoming deeper and more alive as an object of wonder in a person, much as, say, suffering but surviving a dreadful accident might be said to remind one of one's mortality. Thus the connection with sex of reproduction and all it involves casts its shadow over sex in the kind of way that mortality casts its shadow over human life. And this is so even if a given person never thinks of procreation (except, perhaps, to prevent his or her sexual acts leading to conception), just as it is so even if a person never thinks of his own mortality (except to suppress or ridicule the thought). For the kinds of thoughts I have said people have about the birth of a child and those that people have who have survived death form part of the collective experience of mankind, of the wisdom concerning what it is to be a human being and thus of our sense of who and what we are.

I am not claiming, of course, that reflection on the connection of sex with reproduction is the only way in which it is possible for one to come to a deepened understanding of human sexuality. I am just saying that it is a central or permanent way in which this can happen for creatures such as we are, and thus that any account of sexual desire which leaves it out must be inadequate.

As I have already said, however, it does not follow from the fact that sexual desire is capable of deeper forms of expression that only such expression of sexual desire is morally legitimate. Moreover, it is often extremely unclear just which kinds of expression of sexual desire are shallow and which deep. Thus Stefan Zweig, from whose essay on Casanova I have already quoted, manages in that essay to celebrate the very shallowness of Casanova's erotic life, finding in it much to envy in its freedom from moral concerns and in its full-blooded impulsiveness. Yet Zweig would certainly not have supposed it to be good that all behave as Casanova did. It is possible to celebrate the sheer variety of forms of expression of human sexual desire whilst being glad that they remain that, a variety, and that none establishes a hegemony over the others.

Sexual desire, then, I am arguing, is interestingly balanced between depth and shallowness. There is, perhaps, a reason for this in that located close to the centre of our experience of sexual desire is, oddly enough, that of disgust.

In a valuable essay, David Pole has analysed the concept of disgust.[11] He argues, as have others, that disgust always carries a charge of attraction: those things we find disgusting we find both repellent and attractive. Pole also suggests that we get our central notion of disgust from organic matter that is decomposing in some way, which would help explain why such things as slugs - to take one of Pole's examples - are experienced as disgusting: for the slug's slimy body, which it appears to be losing as it crawls along, seems to be caught in a process of decay and corruption. One of the most disgusting things I have ever seen was the neck of an otherwise healthy horse, gashed wide open by barbed wire, into which had buried themselves thousands of maggots which were feeding on the blood oozing in clots from the wound. A friend told me of his disgust on seeing a frog which has a loose back like a string vest into which the young flee to seek shelter and are carried for safety. Organic decay, then, or what looks like it, or smells of it, is perhaps the core of disgust.

Consider now the sexual act. In this act the bodies of those involved undergo profound changes: the flushing of the face, the erection of the penis, the tumescence of the nipples, the secretions of the vagina. One is overwhelmed in desire by one's body, as Sartre puts it: one's will is here in abeyance. All of these things can, of course, be received as an expression of excitement. But there is no doubt that they can be seen as disgusting, and often have been so seen: I should imagine that Christianity has been particularly good at finding them disgusting. For, by their very nature, and in their triumph over the will, they are redolent of a body in decay. This is why desire for the other in his or her flesh can so easily, in certain persons, tip over into disgust with his or her flesh. And in sexual jealousy such disgust is to the fore: for the sexually exciting transformations of the beloved's body resemble nothing so much as the disgusting decay of that body when they are provoked by, and express desire for, a rival. Yet the transformations of one's beloved's body, even when they are connected with one's rival, remain exciting, and they do so even partly because they disgust, for that which is disgusting is appealing, as we have already noted. Disgust, one might say, adapting a Sartrean idiom from another context, lies coiled like a worm at the heart of desire, and it is brought to the light of day by betrayal. Sexual jealousy may begin in the recognition of one's dispensability as a sexual partner, but once it has been evoked it feeds upon the primordial disgust which lies hidden in all sexual acts.

It might be said that the idea that disgust lies at the heart of sexual desire is absurd. And it is, of course, true that not everyone will be susceptible to the sense that the transformations of the body in sexual excitement are redolent of a body in decay, however latent this might be. But there are other reasons for supposing that disgust is inherent in sexual desire. For example, it just seems to be the case that sexual desire (especially male desire?) is often ignited and intensified by a sense of doing something which involves disgust. This is connected with the fact that in sex we suspend or overcome our normal sense of disgust. As William Ian Miller says:

> [S]exual desire depends on the idea of a prohibited domain of the disgusting. A person's tongue in your mouth could be experienced as a pleasure or as a most repulsive and nauseating intrusion depending on the state of relations that exist or are being negotiated between you and the person. But someone else's tongue in your mouth can be a sign of intimacy *because it* can also be a disgusting assault.[12]

But can it be right to say that modern sexual desire, whose expression is so free in comparison with that of previous ages, carries a sense of disgust at its core? Perhaps the idea is not as absurd as it might seem, for A. Béjin has argued that

> present day [sexual] norms tend to provoke a conflict between immediate surrender to the demands of the senses, and an increased conscious mastery of the organic processes... One must... abandon oneself to sensation, without ceasing to submit one's actions to a rational calculation of 'sexual expedience'.[13]

The claim is that we have done a great deal to subsume our sexual practices under the same kind of cost-benefit calculus that applies in so many other areas of our life. If this is right, then

modern sexual desire, for all its seeming liberation from older forms of control, may be thought to express a powerful asceticism which itself testifies to a sense of disgust with sex. Indeed, the fact that modern people seem obsessed with a kind of highly stylised, more or less 'pornographized' sex is itself perhaps a sign of a kind of unacknowledged disgust for sex, a disgust for sex that cannot be packaged and presented in a highly sanitized form.

I have spoken, then, of the possibility of a deepened understanding of sex and of the disgust which is implicit in sex. These two ways of thinking can certainly pull us in different directions, making us think of sex as now something full of grace and light, now as something mean and shabby. But they can pull in the same direction. For the experience of sex can be deeply consoling. If we ask why this is so, then a key part of the answer is surely that, given the wretchedness of the human heart and its potential to fill one with disgust, it can seem little short of a miracle that one person should consent to the intimacy with another that making love involves. In other words, in some moods it can seem that when two people make love this act will depend upon, and involve, mutual forgiveness.

Responding to such a thought, some have seen in sex the possibility of a quasi-religious act, as John Berryman suggests in one of his poems: 'Our Sunday morning when dawn-priests were applying/Wafer and wine to the human wound, we laid/Ourselves to cure ourselves down...' Such an idea is certainly blasphemous, but it helps us see that, in an age of decay of religious belief, there may lie secretly in the modern obsession with sex something more than I have already suggested: a kind of longing for a redemption no longer available in traditional terms.

There is, for some people, something melancholy in the fact that sex can be both a source of the kind of consolation I have mentioned, as well as being imperious and desperate in the way I have also mentioned. We often long for it to express only the most tender of feelings. Yet one can also be glad of this discrepancy in our experience of what sex is, since it makes of sex one of those mysteries of the human condition which help us hold on to the sense that life is worth living because what it offers us is inexhaustibly rich and varied.

King's College London

1 Jean-Paul Sartre, *Being and Nothingness,* tr. Hazel Barnes, (London: Methuen, 1984 [1943]), II, iii, 2.

2 Sartre, *Being and Nothingness,* p.389.

3 Roger Scruton, *Sexual Desire* (London: Weidenfeld and Nicolson, 1986), ch.4.

4 Scruton, *Sexual Desire,* pp.89-90.

5 Stefan Zweig, *Casanova: a Study in Self-Portraiture*, tr. Eden and Cedar Paul, (London: Pushkin Press, 1998 [1928]), pp.88-9.

6 Scruton, *Sexual Desire*, p. 76.

7 Scruton, *Sexual Desire*, pp. 362-3. Cf. also Igor Primoratz' discussion of Scruton's views in *Ethics and Sex* (London: Routledge, 1999), ch.3.

8 Primoratz, *Ethics and Sex,* p.46.

9 Primoratz, *Ethics and Sex,* p.49.

10 D. H. Lawrence, 'A Propos of "Lady Chatterley's Lover"', in *Phoenix II,* Warren Roberts and Harry T. Moore (eds) (New York: Viking Press, 1970), p.504.

11 David Pole, 'Disgust and Other Forms of Aversion', in *Aesthetics, Form and Emotion,* George Roberts (ed.), (London: Duckworth, 1983).

12 William Ian Miller, *An Anatomy of Disgust* (Cambridge, Mass.: Harvard University Press, 1997), p.137.

13 A. Béjin 'The Influence of the Sexologists', in P. Ariés and A. Béjin (eds), *Western Sexuality: Practice and Precepts in Past and Present Times* (Oxford: Basil Blackwell, 1985)

Mat Carmody

The Problems of Vagueness

Introduction

It has just turned seven o'clock on a cold and wet November morning. You know it has just turned seven o'clock because your alarm clock has sounded. You have to get up to get ready for college because you can't pretend to have fallen ill so soon after your last bout of sickness. You find yourself thinking that, what with it being only seven o'clock, there's plenty of time and that it can't hurt to spend just another couple of minutes in bed. How can two paltry minutes matter? Who could deprive anyone of little more than a hundred seconds of extra warmth and comfort on a morning like this?

I'd be prepared to bet that you've reasoned in this way at least once or twice in your life. We are all very willing to tell ourselves that an extra minute won't matter, that an extra few pence is of no importance and that an extra few chips will make no difference to our waistlines. We know that people have been thinking in this way for a very long time. The biblical patriarch Abraham tries to use the insignificance of small differences when arguing with nothing less than God. Yet it is with an ancient Greek philosopher called Eubulides that many people find the start of a deep philosophical puzzle that such reasoning can produce. Eubulides was famous for having seven riddles, which sound like brain-teasers, but which when analysed prove to be of more than coffee-break interest. Two of these are known as the paradox of the heap and the paradox of the bald man. We shall take a look at the first and return to the second in a moment.

The Paradox Of The Heap

I hope you will agree that if I arrange them in the right way, a million and one grains of sand can make up a heap of grains of sand.[1] I hope you will also agree that a single grain of sand does not make up a heap. I put a heap-shaped one million and one grain before you and a single grain behind you. Your task is to remove a grain at random and throw it onto the sand behind you. You are to repeat the task one million times until you have a single grain before you.

It's tempting to think that the addition or subtraction of a single grain can't make a difference to whether something is a heap or not in the same way that the addition of just two minutes to your morning lie-in doesn't really matter. Heaps and lie-ins, we may say, *tolerate* small alterations. We shall call this tempting thought the *tolerance thought.*

Let's now present matters formally. An argument is a collection of sentences we call the premises alongside a sentence we call the conclusion. The premises we assume are true. An argument is a good or valid argument if there is a logical path from the premises to the conclusion. Let's put this thought and the two we opened with as the premises of an argument.

(1) 1,000,001 grains make up a heap

(2) 1 grain does not make up a heap

(3) The removal of 1 grain does not turn a heap into something that is not a heap (or vice versa).

What conclusion can we draw? Each time you remove a grain from the heap, you preserve the heap because of (3). Each time you add a grain to the 'non-heap' collection behind you, you still have a non-heap, also because of (3). Even after a million subtractions to the heap before you therefore, when just one grain remains, you still have a heap:

(4) 1 grain makes up a heap.

Similarly, after a million additions to the non-heap behind you, you have a 1,000,001 non-heap:

(5) 1,000,001 grains don't make up a heap.

We can therefore deduce, by combining (2) with (4) and (1) with (5):

(6) 1 grain makes up a heap and 1 grain doesn't make up a heap.

(7) 1,000,001 grains make up a heap and 1,000,001 grains don't make up a heap.

Two sentences present a contradiction if they both cannot be true together and they both cannot be false together. For example, the sentences 'John is at home' and 'John is not at home' present a contradiction, if we understand them literally as talking about the same John. A paradox is an argument whose conclusion presents a contradiction. (6) and (7) both present contradictions. Our argument is a paradox: the paradox of the heap.

You should be able to see that the same argument can prove that 2 grains do and do not make up a heap, that 3 grains do and do not...that any number indeed do and do not. This is clearly unacceptable.

Our conclusion is the product of the premises and the logical reasoning we used to move forward from them. If we can't accept the conclusion of an argument, it is with either or both of these factors that we must find fault.

Reasoning

It often happens that we are tripped up by an argument that looks and feels logically valid but which turns out to be more complex than we think and conceals logical blunders. Yet this argument really just uses one logical principle.

The first principle we rely on is called modus ponendo ponens or (usually) just modus ponens for short. The principle says:

(MP) From [A] and [if A then B] deduce [B]

For example: from [it is raining] and [if it is raining, then Bernard will be at home] we can deduce [Bernard will be at home].

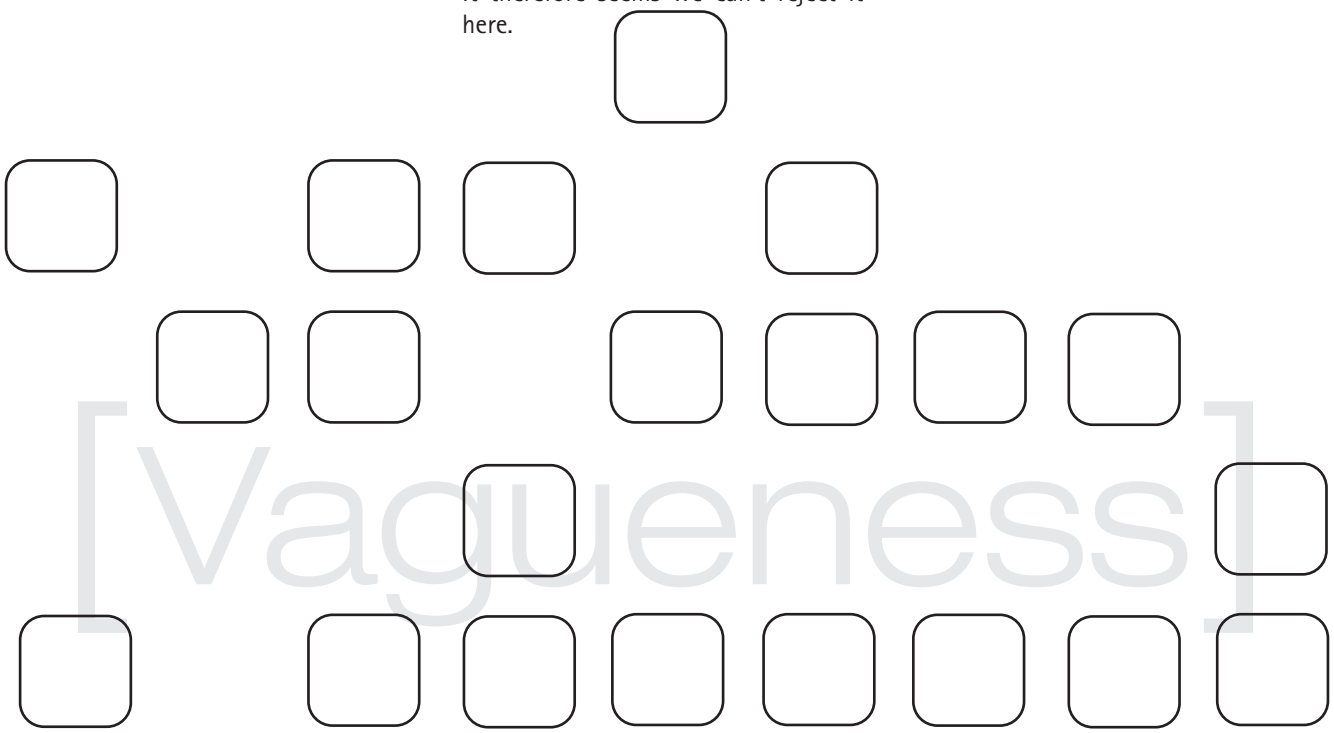
Modus ponens is the cornerstone of reasoning. It captures the idea of conditional judgement or reasoning or action, without which neither we, nor any other creature capable of the most basic of thoughts, could do anything. It therefore seems we can't reject it here.

Modus ponens allows us to deduce [1,000,000 grains of sand make up a heap] from [If 1,000,001 grains make up a heap then 1,000,000 grains make up a heap] and [1,000,001 grains make up a heap]. The conditional there is just an instance of our tolerance thought. By repeating the reasoning, we reach (4): 1 grain makes up a heap.

(2) and (4) present a contradiction. Formally, we can join them together to come up with the sentence: 1 grain of sand doesn't make up a heap and 1 grain of sand makes up a heap. In so doing, we use another logical principle, called *conjunction introduction.* A conjunction is just a sentence of the form [A and B]. The principle says:

(Conj I) From [A] and [B], deduce [A and B]

This should seem too obvious to merit either denial or comment. We shall therefore conclude that the reasoning is not at fault.



Nihilism

It must therefore be one or more of the premises. When we are faced with premises all of which seem true, we need to think of what it would mean if they each weren't true and which of these situations would be the least unacceptable. If (1) were not true, then 1,000,001 grains would not make up a heap. I chose this number thinking that that many grains could be formed into a heap. I hope you'll think that, if I am wrong, then some large number of grains could be made into a heap. If we deny (1), therefore, we are in effect denying that any number of grains could make up a heap - or, more simply, that there are heaps. Similarly, if (2) were not true, we'd be in effect denying that any number of grains could fail to make up a heap.

Either denial seems impossible. It is however as close to an axiom in philosophy as anything is that there is no position on a problem, however crazy, that no-one has not chosen to adopt. It so happens that a couple of philosophers have defended the thought that (1) is false. The idea is that the paradox shows our vague words to be so defective that they don't truly apply to anything. Nor could they apply to any possible thing. It is rather as if 'heap' simply isn't fit to be applied to things in the same way that 'red' isn't fit to be applied numbers. In saying that 'three isn't red', we aren't saying that it is some other colour. Numbers don't have colours. Similarly, collections of material objects can't be made into heaps.

Most people regard this view as both nonsensical and desperate. Even if we allowed that nothing really is a heap, then plenty of things *look like heaps or are heaps according to how we use that word.* You should be able to see that these expressions are just as vague as 'heap' is. It therefore follows that nothing even looks like a heap. Now we may be wrong about whether someone is tall because appearances are misleading. Yet whether someone looks tall just depends on appearances - indeed, how things appear to us individually. Generalising, if we allow that 'looks X' doesn't apply to anything, where X is replaced with any vague expression, it seems that we can't even describe how things seem. Yet we can surely do this at least. Even Descartes allowed that we couldn't be fooled about our most basic appearances.

Almost all philosophers think that the problem lies with the third premise. In order to explain why, it'll be worth looking at the puzzle from a different angle.

Vagueness, Boundaries, and Borerline Cases

It is often said that not everything is black and white. An action may not be clearly good or clearly bad but be good in some respects and bad in others. An accusation may neither be wholly true nor wholly false. An expression may neither clearly apply nor clearly fail to apply. For example, you can probably think of someone who you wouldn't call tall but of whom you couldn't really say that he isn't tall. He's somewhere in the middle. He's what we call a *borderline* case of tall.

If you took someone who was tall and gradually shrank them, then you'd find yourself with a borderline case of tall before finding yourself sometime later with someone not tall. Similarly, if you took something that was a heap and removed grains one at a time, you'd be faced with borderline cases of heaps before non-heaps. *Heap* and *tall* are alike in being *vague.* They allow for borderline cases. They invite the tolerance thought. Although we haven't formally spelled out the paradox for *tall,* it is not hard to do. Do you think a difference of a millimetre matters to whether someone is tall or not? If not, then by shrinking a person two metres in height by a millimetre at a time, you can end up with a person one metre tall that is still tall.

An expression such as *eighteen years of age or more* is one that we can regard as *sharp.* You are either eighteen years of age or more or you are not. The age of eighteen is a sharp boundary that separates these two possibilities. The thought is that *heap* and *tall* are vague just because there is no such sharp boundary. If there were a sharp boundary for *tall,* then there would be some exact height above which people were tall. If there were, then someone a millimetre below this height and therefore not tall would become tall by growing a millimetre. This contradicts the tolerance thought. The fact that we are tempted by tolerance shows that we don't like this idea of a sharp boundary. Indeed, we think that there are instead borderline cases, which replace this sharp line with a middle region.

Let us now return to the third premise. It says that a small change of a grain can't turn a heap into a non-heap. If this were false, wouldn't it mean that a small change *could* turn a heap into a non-heap? It seems not. This would only be so if we think that there were two possibilities: heap and non-heap. No small change can turn a heap into a non-heap just because there are lots of borderline heaps in between. We can therefore say that the third premise isn't true without this committing us to a heap/non-heap sharp boundary.

What Vagueness Is and Isn't

Any expression that allows for borderline cases and the tolerance thought is vague. We've seen that *heap* and *tall* are vague. Eubulides invited also to think that by pulling one hair out of the head of a non-bald man, you couldn't make him bald. You should be able to construct a paradox using this tolerance thought along the above lines. Such paradoxes are called *sorites paradoxes* after the greek word for 'heap'.[2]

The ingredients needed to generate these puzzles are simple. You need to think of a dimension of change, such as numbers of grains or numbers of hairs. At one end of this dimension, we find things that are X. At the other, we find things that are NOT-X. All we add now is the tolerance thought: that small differences don't matter.

Here are some examples to get you thinking. Heights, weights, distances, times, temperatures are all dimensions of change. If a person is two metres in height they are tall and if one metre they are not tall. Yet can growing by a millimetre make someone tall who was not-tall? You should be able to generate similar puzzles for *short, heavy, light, near, far, hot, warm and cold.* Now think of any expression for which number is relevant. Alongside heap and *bald* we have *crowd* (number of people), *fleet* (number of ships) and *many, few, lots, some, plenty* (number of whatever is counted). Time is also a dimension of change.

It seems odd to say that the passing of a second can make something *old* that wasn't before, along with making a creature an *adult* of its species if it wasn't before. Finally, take any everyday material object, such as a *coffee cup*. It is made up of a large number of molecules. If I remove a very small number at random, I surely can't destroy the object by turning it into something no longer a *coffee cup*. The same is therefore true of any of the things that populate our environment - including ourselves. If *you* can tolerate the loss of a few molecules, then you do and don't exist!

Philosophers distinguish different types of vagueness. One type with which sorites vagueness is often confused is what I shall call 'comparison-class vagueness'. Suppose I tell you something has a height of ten metres and I ask you whether it is tall or not. You can't answer until you know what sort of thing it is. A ten-metre tall man is very tall. A ten-metre building is not tall. A ten-metre tree may be tall, depending on the species. It can therefore be vague whether something is tall just because it may be vague what sort of thing we're using as the basis for making our comparative judgements. This sort of vagueness is not at all the same thing as sorites vagueness. Once we have agreed that the thing is a beech tree and that it is indeed tall, we still generate a paradox for 'tall beech tree' via the tolerance thought that if a tall beech tree shrinks by a millimetre, it is still tall.

A third and similar form of vagueness is informational vagueness. You want to meet Anna for lunch. You ask me where she is. I tell you that she is in London. This information is very likely to be too vague to be of much use. What you wanted was her precise location. On the other hand, if you know that Anna only ever has lunch in one restaurant when she is in London, then my telling you that is in London rather than Paris will suffice. Information is vague or precise dependent on the use we intend to put it to.

Three Truth-Values

The sentence 'Jupiter is a bigger planet than Mars' is true. Another, more formal way of putting it, is that the sentence has the truth-value *true*. In the same way, we can say that the sentence 'Mars is a bigger planet than Saturn' has the truth-value *false*.

So far, we have spoken as if these are the only two values a sentence can have. We believe that it is true that 1,000,001 grains make a heap. What about 1,000,000 grains? Either the sentence '1,000,000 grains make a heap' is true or false. If false, then a sharp boundary divides these two collections into *heap* and *non-heap*. Our tolerance thought says that the sentence is true, just because one grain can't make a difference. But this generates a paradoxical conclusion.

We recently found a way through the dilemma: we admit borderline cases. At the level of truth, this translates into adding a third truth-value: *borderline* or *indefinite*. It may then be, for example, that '34,346 grains make a heap' is *indefinite* just because they make up a borderline heap.

If we turn back to the formal argument, we are saying that (1) and (2) are true and that (3) isn't. Should we say that (3) is indefinite or false?

(3) The removal of 1 grain does not turn a heap into something that is not a heap

(or vice versa).

One generalised and clear way of re-expressing (3) is as follows:

(3') If a collection of x grains makes up a heap, then a collection of x-1 grains makes up a heap.

We might think as follows. (3)' says something untrue, that is, false, namely that 'heap' continues to apply, no matter how few grains remain. This doesn't mean that we have a sharp borderline. We interpose the borderline cases.

Unfortunately, we can't say that - at least, not so quickly. It is not possible to explain here in detail why. You will remember that earlier we looked at two principles (MP) and (Conj I). These seem to capture some of the meaning of our everyday logical words 'if' and 'and'. It happens that, if we say that (3) or (3') must be false, we end up forced to challenge the logical principles we hold for these words and others like 'or', 'not', 'all' and 'some'. It turns out that making (3) or (3') false is too high a price to pay.

So we must hold that (3) and (3') are indefinite. What does this mean? One thought would be: it's indefinite because it depends on the number of grains. But this still implies that, for at least one x, if a collection of x grains makes up a heap, then a collection of x-1 grains makes up a heap, which we don't want. A second thought would be: it's indefinite because we can't tell. But this makes the problem one of knowledge: it may be that the removal of a grain can turn a heap into a non-heap but we can't be aware of this. This isn't any more tempting.

We can in fact defend the thought that it is indefinite from a logical point of view and that this doesn't translate into anything 'intuitively' obvious in everyday English. Yet you might think that (3) and (3') are better described as not indefinite, not true nor false, but nearly true. Can we make sense of this?

An Infinity of Truth-Values

Some philosophers think that there are far more than three truth-values. Alan and Bill are 1.67m and 1.68m tall respectively. They are both borderline tall. It is indefinite whether they are tall or not. Yet Bill is taller than Alan by a small amount and so it should be possible to say that it is truer that Bill is tall than Alan is tall.

We can distinguish many degrees of tallness, of baldness, of redness, and so on. Indeed, every vague expression X generates a meaningful expression '...is more/less X than...'. We should therefore have as many degrees of truth as we can have degrees to which these expressions apply. Between the numbers 1.67 and 1.68 are infinitely many numbers. We should therefore admit infinitely many truth-values. Even if there aren't infinitely many degrees of being a heap, it will be better to have too many truth-values which we can group together than to have too few.

As far as the paradox is concerned, a very neat reply becomes possible. I shall leave out the technical details and try to convey the essence. It is true that 1,000,001 grains make a heap. It is slightly less true that 1,000,000 grains make a heap. It is slightly less true again that 999,999 grains make a heap. It is false that 1

grain makes a heap. Our premise (3) tells us that it is true that if x has n grains and is a heap, then x' with n-1 is a heap. We can say that this is nearly true but not quite true. What is quite true is that if if x has n grains and is a heap, then x' with n-1 is a heap to a slightly lower degree. We confuse these two thoughts. If we admit that (3) is almost but not quite true, it becomes possible once again to defuse the paradox. This is better than saying it is simply 'indefinite'.

Epistemicism

Those who defend alternative logics consider vagueness to be a semantic problem. They think that words like *heap* and *tall* have meanings that are incomplete. *Heap* is defined in such a way that some things fall under the definition - the clear cases of heaps. The definition equally excludes some things - the clear cases of non-heaps. The borderline status of a borderline case reflects a real and profound absence of truth or falsity stemming from the definition not covering such cases.

An alternative and radical proposal is that vagueness does not arise from incomplete meanings but from our knowledge of meaning. It is not semantic (to do with meaning) but epistemic (to do with knowledge). The Epistemicist tells us that *every* vague expression is semantically precise. In other words, there is a sharp boundary dividing heaps from non-heaps, tall people from non-tall people, and so on. We simply don't know where this boundary is. Furthermore, we can't know.

In order to explain why, let us examine the concept of knowledge. We say things like 'Bernard knows that we are meeting this evening at 8pm'. We can break this down into three parts: a subject (Bernard), a *propositional attitude* (knowing that) and a proposition (we are meeting this evening). For our purposes, we can think of a proposition as like a sentence about which it makes sense to ask: is it true or false? By way of comparison, if I say 'Bernard knows how to drive a car', we have an attitude - knowing how - but not a propositional one, as it does not make sense to ask whether 'to drive a car' is true or false.

Our 'attitude' to the proposition 'we are meeting this evening' may be one of knowledge but also one of belief, desire, fear or hope, amongst others. An attitude is something like a way a mind relates to information. Alongside knowledge, we have the closely-related attitude of belief. The essential differences between belief and knowledge are as follows. Firstly, you can believe something false but you can't know something false. It is possible to say 'Until I was eight, I believed New York was the capital of America' but not 'Until I was eight, I knew New York was the capital of America'. Secondly, knowledge is more 'robust' or reliable than a true belief. I ask you whether Bernard was at home yesterday and you reply, 'I believe he was'. If I ask, 'are you sure?' it would be fair for you to say, 'no'. You have some grounds to believe he was at home but you may not be supremely confident. If you reply, 'I know he was', then you advertise your confidence. When we want to find things out, we tend to search for people who know, rather than simply believe.

Consider now the following situation. I am at a football stadium in which there are 74,362 people, including me. I guess how many people there are and I guess that there are 74,362. Now you will agree that, even though I am right, I don't know that there are 74,362. I am right by chance. I haven't done anything, such as count the people, to make my lucky guess into a robust piece of knowledge. Indeed, had there been a few hundred more people, I could have just as easily guessed, wrongly, that there were 74,362 people because it would have looked no different to me. I may not be able to have exact knowledge of the number of people in the stadium but I can have inexact knowledge. I can know that there are at least 2 and not more than 1,000,000 people in the stadium just by looking because these two situations look sufficiently different from the actual one for someone with my capacities for discrimination. If I had a better capacity to judge crowd sizes, I would have less inexact knowledge. I might know, for example, that there are between 60,000 and 80,000 people.

The phenomenon of inexact knowledge is very general. You can know on the basis of appearances someone's height or age or weight or distance from you roughly, which is to say that you can say what heights, ages, weights and distances are clearly wrong and thus define a range of possible answers in the middle. The better a judge of stadium capacities, heights, ages, and so on, the smaller the middle range wherein lies the right but inaccessible answer.

The Epistemicist likens the exact number of people in the crowd to the exact number of grains he says is the sharp boundary between a heap and a non-heap. Let us call that exact number n. As a speaker of English, I understand the word 'heap' and, when presented with heaps, borderline heaps and non-heaps, will react appropriately. As with so many of our words, our understanding consists in being able to use 'heap' in a certain way rather than being able to voice a clear definition of heap. (Look around you and find a word to classify each object you can see. Can you define any of these words precisely? Do you feel this undermines your claim to be using these words correctly?)

The Epistemicist says that how we use 'heap' in fact determines n but that we don't know how our collective use fixes n. It is not as if we can examine every possible situation each of us could be in to see how each of use would use 'heap'. Since we can't see this 'total use', we couldn't discriminate it from the total use of a group of speakers of a language Schminglish who used 'heap' very much like us but not identically to us. Since it is use that determines the sharp boundary and the uses are very similar but distinct, 'heap' in Schminglish determines a sharp boundary of - let us say - n+1 grains.

By way of analogy, think of the speakers of English as a stadium with 74,362 people and the speakers of Schminglish as a stadium with 74,363 people. In the same way that we can't distinguish their total number in either case, making them appear indiscriminable, we can't distinguish how English and Schminglish speakers use 'heap'. So since we can't distinguish our use precisely, we can't distinguish the number n precisely. We can only have inexact knowledge. Just as I can know that the stadium has more than 2 people and fewer than 1,000,000, I can know that 2 grains can't make a heap and 1,000,000 can. Just as I can't know that the stadium contains 74,362 people, I can't know whether 74,362 grains make up a heap. It is a borderline case not because there is no right answer, as the semantic approaches claim, but because I can't know the right answer.

So Who's Right? Logic vs. The Incredulous Stare

The semantic approaches start from the simple observation that the sorites paradox supposes that every

statement is true or false. This supposition is called the 'principle of bivalence'. The common recommendation is then to reject this principle and allow for more truth-values. The common problem is that the logics that result don't seem to work properly.

We saw earlier two logical rules (MP) and (Conj I). A 'logic' is a collection of such rules which tells you what you can and cannot deduce from your premises. In order to design a set of rules, you have to decide how many truth-values you are going to allow. For you are in effecting asking yourself: if A is true/indefinite/etc. and B is true/indefinite/etc., then can I deduce C? Classical logic is a system of rules with the principle that there are two truth-values, **true** and **false**. Non-classical logics are those that have different sets of rules and/or more truth-values.

Classical logic is highly regarded because its rules capture what we feel to be patterns of proper reasoning. It is therefore claimed by many to be logic that not merely our minds, but any intelligent mind, would employ.[3] It is therefore claimed that, because of this, we should accept that every sentence is indeed either true or false. This claim would be undermined, of course, if non-classical logics outperformed classical logic when it came to representing reasoning. The problem is that each non-classical logic comes in for heavy criticism on this score. It is alleged that they each have rules that permit deductions that are unacceptable. It is not possible to go into details here. I shall just say that one of the key planks in the defence of the Epistemic position is a demonstration of just how hopeless the opposing positions are on the matter of logic.

The Epistemic position needs to build a solid defence of itself because its central claim is *prima facie* so outrageous as to lead one to think that *any* other position is preferable. Its central claim is that every vague expression does in fact have a sharp boundary. Yet there is something disturbing about the thought that something we do determines a sharp boundary that we can't find. We should be happy to accept that there are scientific and mathematical facts that may escape our knowledge because we lack the equipment to measure them and the minds to understand them. But can it be true that there are facts about who is tall and who is bald that are beyond us?

Each position has its advantages and drawbacks and sorites paradox remains unsolved. Let's finish by looking at two issues that follow from the initial puzzle and illustrate the breadth of the issues vagueness raises.

Higher-Order Vagueness

The Epistemicist asks us to believe in simplicity. There is a sharp but hidden boundary dividing the Xs from the non-Xs for all vague expressions X. The Semanticist rejects this. But is he merely shifting the problem?

One semantic approach said that there were three truth-values. Things are divided into X, borderline case of X and not-X. If so, does this not mean there is not one but *two* sharp divisions? Consider a line of people who differ gradually in height. The tallest is 2m tall, the smallest 1m50cm and each adjacent pair differ by a half-centimetre. We don't believe that there is a sharp boundary dividing them into tall and not-tall. Is it any more obvious where the boundary between the tall and the borderline tall is?

Our second semantic approach said that there were lots of truth-values. Things are divided into many different degrees of X, from 'full' to 'zero'. If so, does this not mean that there are lots and lots of sharp divisions? Here, one might just say: yes. Every difference in height is reflected by a difference in degree of being tall. We therefore argue as follows. A man 2m10cm in height and a man 2m in height are both simply tall. *Tall* covers a range of heights. The same is true of not-tall. In between those heights tall and not-tall are lots of heights and lots of different degrees of being tall. Let's lump them all together as reflecting different ways of being borderline tall. In this way, we resurrect the problem of the previous paragraph. It seems that at some point, we must cross a height marking the end of tall and the beginning of borderline tall. But where?

It has seemed to many philosophers that the problem doesn't arise because of *higher-order vagueness.* So far, we have said it can be vague whether someone is tall or not tall and that such a someone is borderline tall. But couldn't someone be a borderline case of borderline tall? If so, then there is no boundary between tall and borderline tall. There is a range of heights corresponding to borderline borderline tall.

You may have already guessed what the next problem is going to be. If borderline borderline tall is sandwiched between tall and borderline tall, then we have *even more* sharp boundaries. Where does tall cross over into borderline borderline tall? Of course, if borderline borderline tall can itself have borderline cases, then we can squeeze this new category into where a borderline was supposed to be. But the problem simply reappears at the next level up.

Many philosophers think that we have to admit these higher levels or orders of vagueness just because it is absurd to suppose that there are sharp boundaries between tall, borderline tall and not-tall. Yet it is not obvious that it is a price worth paying. It seems that if we start 'going up' we must go up indefinitely. This means that tall makes infinitely many fine divisions. Can this simple little word like *heap* hide so much complexity?

A plainer line of attack is that it is entirely spurious, at least as a means of avoiding sharp boundaries. We reason just as we did a moment ago. There are some things that are, simply, tall and some that are, simply not-tall. In between, let there be as many divisions into degrees of tall and ever-higher orders of borderline cases of tall as you wish. Collect them altogether and label them as n = 'not clear cases of tall and not-tall'. Once again, we seem to have arrived back with three categories: tall, n, and not-tall. Once again, we seem to have arrived back with two sharp divisions, marking the edge of 'tall' (with everything else) and 'not-tall' (with everything else).

We don't appear to be any more knowledgeable about these boundaries. So where epistemicism asks us to believe in one hidden and inaccessible boundary, all semantic approaches ask us to believe in (at least) two hidden and inaccessible boundaries. It is this strange fact that Epistemicism asks us to believe that provides a lot of the drive to find an alternative. If semantic alternatives are in the same boat, however, then Epistemicism comes off in a better light. Or, you might conclude, no-one comes off in a good light at all. Vagueness seems to commit us to sharpness wherever we turn! There is therefore something very deeply wrong in how we understand the relation between minds, language and the world.

Onticism

Some say that vagueness is a problem with defective meanings. Others say that it is a problem of knowledge of those meanings. Others still say that vagueness is a feature of things.

This is a natural view when we are talking about things we can refer to with names, such as objects and places. Does a cloud occupy a precise space or is it a vague entity with fuzzy spatio-temporal boundaries? Does London have a sharp geographical boundary? Or is it vague whether some bits of land are part of London or not? Does London have a sharp chronological boundary? Did it come into being at some precise moment? Or was the transition from the first settlement to the final city a vague one?

Many philosophers think that reality itself is not vague. A man has a precise height but may be vaguely tall because of a problem with the word *tall.* Equally, there is a precise configuration of buildings and bits of land in the world but which parts fall under 'London' is vague through it being unclear what that word means as well.

One reason they think this is because of a famous little argument that appears to prove that it is incoherent to suggest that the world might really be vague. It goes as follows. Let us suppose it is vague whether London today ($London_{2004}$) is identical with London a thousand years ago ($London_{1004}$). It is surely not vague

that $London_{2004}$ is identical with $London_{2004}$ - that's trivial! So there is something $London_{2004}$ has that $London_{1004}$ lacks. This is the property of being non-vaguely identical to $London_{2004}$.

There is a widely-held view that if x and y are identical, then they have identical properties. It follows that if x and y do not have identical properties, they are not identical. We have just found a property that $London_{2004}$ has that $London_{1004}$ lacks. So they are not identical after all. This contradicts our assumption that it was vague whether they were identical. The idea that two things may be vaguely identical is therefore paradox-inducing and unacceptable.

If objects can be vague, then it will not always be clear whether one object is identical to another. So if x is a vague object, there will be an object y such that it is vague whether x = y. We have just proved that it is not possible to say that it is indefinite whether one object is identical to another. It follows that objects cannot be vague!

This might seem a bit fishy. Does this little piece of reasoning really prove that all clouds are precise entities? According to most philosophers, the answer is no. It is true that clouds are vague entities and it can be true to say that it is indefinite whether cloud x is identical to cloud y, just as common sense demands. What the argument shows is that we can't make sense of this by supposing the vagueness is somehow a feature of the clouds themselves. That leads to paradox. The right conclusion is that vagueness is a feature of how we refer to these things.

The semanticist says that when I use the word 'cloud' in 'that cloud is fluffy', I don't pick out a particular vague entity. I don't manage to pick out any one thing at all. It is vague what particular region of space I am referring to. If I wonder whether cloud x is identical to cloud y, this can be vague just because it is vague which two regions I am thinking about.

The Epistemicist says that I do pick out a precise region of space but I don't know which one. When I wonder whether cloud x is identical to cloud y, then either it is or it isn't. I may not be able to know the right answer. When this happens, it will be vague whether they are identical or not.

Conclusion

The vast majority of our words are vague. We are seduced by the thought that small differences don't matter when it comes to using them. Yet this leads to the unacceptable conclusion that they apply to everything and nothing. The problem of vagueness is to understand how vague expressions do have limits and how small differences can matter. A solution to this will be a solution to the sorites paradox. As with so many of the best philosophical problems, it is simple and fundamental and continues to vex philosophers over two millennia after it was first written down. Not *so* much so, however, as to get them out of bed any earlier on cold winter mornings.

1 We shall assume from now on that we are dealing with heaps of grains of sand. We shall also assume that collections of grains of sand have, where necessary, the right structures and shapes to be heaps.

2 'sorites' translates as 'heaper, one who heaps'

3 We are not trying to capture, in a logic, how we actually reason, because we often reason illogically. We are trying to discover the correct way to reason. In the same way, when we do mathematics, we're not investigating how we actually operate with numbers, because we often make mistakes. We are trying to discover the right answers - the answers we should get to if we reason correctly.

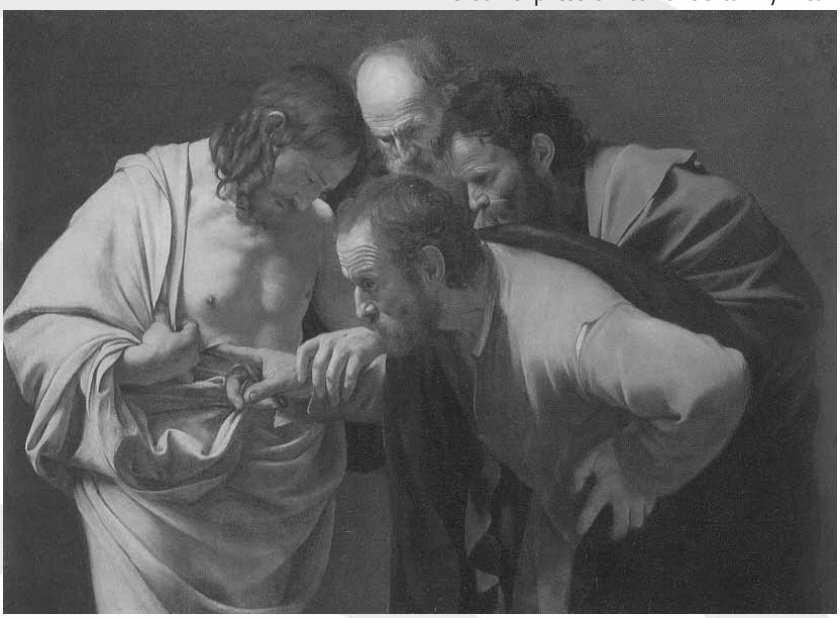
Jennifer Booth

Scientific [Knowledge]

Truth, Induction and Falsification

Prima facie it seems obvious that the findings of scientific research constitute knowledge. Albeit often a specific and highly detailed type of knowledge, but knowledge nonetheless. The question I want to ask concerns what exactly our common sense notion of scientific knowledge amounts to. In doing so, I suggest that in order to understand a conception of what scientific knowledge is, we must look to how we characterise science in itself. The reason being that the model of scientific methodology that we choose to adopt will govern what *kind* of knowledge (if any) science can be seen to produce. What we need to ask is which of the models purported to capture scientific methodology will best account for our common sense view of scientific knowledge. It is in attempting to answer this question that I turn to two contrasting accounts of scientific method: inductivism and falsificationism.

First and foremost however, we must decide on what exactly it is that we commonly understand to be scientific knowledge. One way to do this is by first explaining what types of knowledge do not qualify. There is a type of knowledge best described as 'practical' knowledge, which involves such things as knowing how to ride a bicycle or brush one's teeth comprehensively. Such knowledge is non-conceptual or basic and it can exist without our knowing how to give verbal expression to it. Certainly I can choose to reflect upon my possession of such knowledge, I can state that I *know* how to ride a bicycle, but this reflection is something quite separate to the practical knowledge itself. It seems the case that such pieces of knowledge can act as their own explanation. That is, if someone questioned whether I could ride a bicycle, all I would need do to exhaust his or her query is to demonstrate my ability. There is no 'higher level' explanation that I should give in order to justify my knowledge. This is not the type of knowledge I have in mind for scientific knowledge.



Another type of knowledge which I mean to avoid is that which grounds our 'rule-governed'[1] behaviour. When we operate within our society we do so according to a body of rules or laws that we must abide by, if we are to remain in that society. Such things as 'driving on the left' or 'stopping at a red light' are examples of this kind of knowledge. It is perhaps best called 'procedural knowledge'. Although this knowledge has conceptual content which can be expressed in such propositional form as 'all cars must stop at red lights', it is not the kind of knowledge with which I am concerned. If we are asked to justify or explain a particular instance of this knowledge, we can only do so by referring to concerns outside the actual content of that knowledge. That is, if asked why it is that you stop the car at a red light, it would be

unlikely that the explanation you would give would be one involving certain refracted wavelengths falling on your retina, being interpreted by your visual cortex and resulting in leg contractions which stop the car. Instead, you would offer either some general reason 'that its part of the legal highway code' or some moral concern 'because I don't want to hurt anyone'.

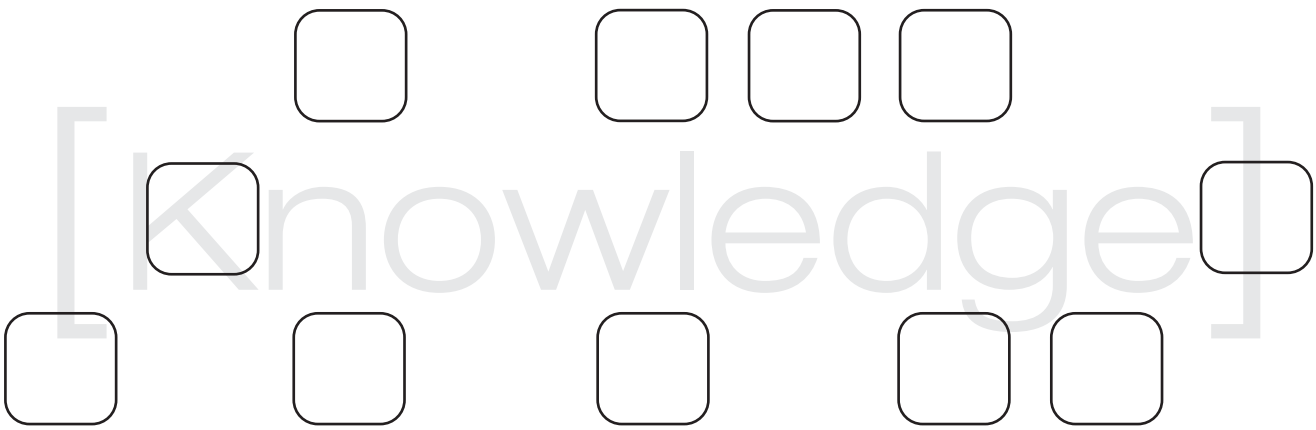
The type of knowledge I have in mind is that which we claim offers us explanation of the way the world is. It is a type of propositional knowledge that takes the form of statements such as 'cats miaow whereas dogs bark', or that 'kicking a football makes it move'. *I will call this knowledge 'operational knowledge' as it seems both a result of, and a basis for, our operation within the world.* It is this type of propositional form that we tend to give to scientific claims about the world. Note that this is not say that all our claims to operational knowledge are obviously reducible to, or dependent upon, scientific claims. On the contrary, scientific claims can be seen on occasion to offer complimentary but quite separate explanations of the world. In terms of the above examples, scientific knowledge might make claims as to the biology of a cat's larynx, and the principles of movement and inertia.

The common view of scientific knowledge is that it constitutes some specialised part of this body of operational knowledge. The kind of claims I have in mind could be such things as 'water boils at 100 degrees centigrade' or statements about the relationship between force, mass and motion. Any reliable law-like proposition that scientific research develops I suggest comes under this category.

So let us return to the key question for scientific methodology, which now looks something like this. In what sense, and to what extent, is science able to provide us with such a specialised subset of operational knowledge? On the one hand, science can be seen to provide us with universal laws, which we put to work in both explaining and predicting our individual observations of the world. On the other hand, the business of science can be seen to lie in the questioning of such laws, in showing how to refute them, and thereby taking our operational knowledge closer to the truth of the world. Put crudely, these two approaches to scientific practice are the methods of inductivism and falsificationism respectively. Let us look at inductivism first.

The methodology of induction in science is simple: we note an event or observation repeated again and again across time and we judge this regularity to be something we can depend upon. Each individual observation reinforces our belief in a characteristic of the world that perseveres beyond a specific time and circumstance. For inductivism, collating observations in science leads to the postulation of various scientific laws. On the basis of these laws, I presume both that an observation I may have has been had before, and that it will continue to be observed on other occasions by other people long after I am gone. In other words, these universal scientific laws offer some kind of certainty to my individual observations; what I see on each occasion is grounded by scientific rules. In short, these laws tell me how the world necessarily is, and so give reason for my trust in this inductive process.

So the basic application of our inductive reasoning is two fold: firstly we think we can explain what we *have* seen by the use of universal laws, and secondly, that we can use these established laws in predicting what we *will* see.



There are however, a number of problems with this picture, concerning both the mechanics of this inductive process and the characterisation it gives of scientific knowledge. First and foremost there is the question of whether we are justified in formulating these universal laws simply on the basis of a discrete number of past observations that we have made.

Taking a striking example, that owing to the scientific observation of planetary motion, science can suggest the guaranteed law that 'the sun will rise every day'. Prima facie, just because the sun has risen in the past, it does not mean that it will continue to do so either tomorrow or the next day. There is, as it were, no guarantee that we will ever see the sun rise again. If this is the case, then it is certainly detrimental to the sense of faith we have in our scientific laws of planetary motions. What we rely upon is the supposition that some kind of necessity has caused the sun to rise in the past and will therefore continue to cause the sun to rise in the future. The problem is that any certainty we think we can obtain from an induced scientific law turns out to be of no more use to us in guaranteeing the truth of the world than any individual observation in itself.

To understand why, let us look to Hume. He stated that when we observe two events to be causally related, say a kick (e_1) causing a ball to move (e_2), what we in fact observe is only a contingent conjunction of two events. That is, the causality that we think we perceive is not actually 'out there in the world' for us to observe. When we see two events and *judge* them to be causally related, it is merely through a habit of the mind, something we project onto the world. A necessary causal link, as such, is not guaranteed. In the relevant parlance, there is no necessity in e_2 following e_1; nothing to guarantee that the next occurrence of e_1 will be followed by e_2.

If we accept this proposal, it is a short step to Hume's general attack on inductive reasoning: that there is nothing in our observation of past or current events that can tell us about the way the future will be.[2] If we presume otherwise, we are begging the question of the uniformity of nature: that what has always been will (for apparently no good reason) continue to be.

If we return to our example of the sun rising every day and the scientific explanation we propose to give in terms of the causal effects of planetary motions, we can see that the Humean objections take hold in the following way. Whatever scientific explanation I give in terms of the observation of regularities in planetary behaviour, no number of observations gives me the right to postulate a universal law. There is no in-built necessity which we can observe that tells us the planets will always move in such a fashion, and that the sun will thus appear to rise every day. We simply cannot postulate universal laws that tell us the way the world irrefutably is and will always be unless we have reason to trust such generalisations.

Besides which, even if we could trust such universal laws as 'the sun will always rise' it is not even clear how many times we would need to see the sun rise in order to justify first proposing this law. That is, how many 'repeated observations' will suffice for us to be certain that something is going to continue to be the case, either in everyday life or the laboratory? It seems that in this case scientific explanation, although detailed and informative, has no claim to being the irrefutable truth of the matter. Indeed, if scientific method is restricted to induction, it seems our claims to operational knowledge are not as certain as we like to think they are.

These objections give to induction a kind of naïve quality, it being merely a habit of the mind to which we succumb. A method which promises universal pieces of knowledge, but inevitably leaves us open to unwelcome error. If we can do no better than to follow this inductive rationale then we are as naïve as Russell's chicken: fed every day by the farmer, expectant of food in the future, and genuinely surprised by the unforeseen wringing of its neck.[3]

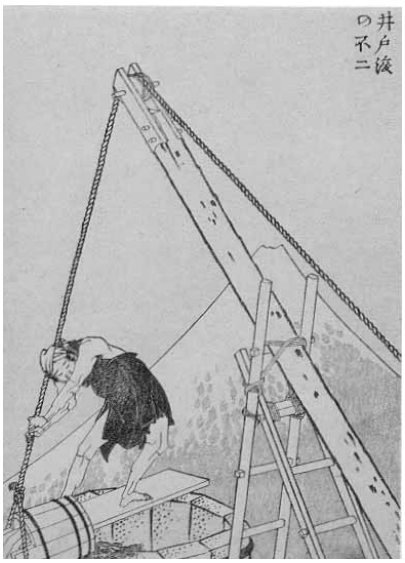
So much for induction as providing a basis for this scientific operational knowledge. As it stands, this methodology seems far too error prone and dogmatic to be the correct explanation of how we understand the characteristics of the world.

Karl Popper claimed that this naïveté which induction displays is a result of its non-rational or non-logical make-up. Although not irrational, the practice of induction does go beyond what is strictly logical. That is, proposing an 'all encompassing' law on the basis of a finite number of observations is thereby to propose knowledge of that which we simply don't know. If scientific method is to be rational, according to Popper, it must only make claims to knowledge that are logically sound. Science is therefore not in the business of making grand universal laws, but instead should concern itself with the examination of individual observations.

For our present concerns, Popper's

model of falsification can be seen in the following way: when we make claims to operational knowledge, we propose certain non-logical 'inductive' laws, and we then rely on science for their *evaluation.* That is, the role of science is different to that proposed by the method of induction. The job of science, according to Popper, is to take these laws which we make through habit or generalisation, and try to prove them wrong. That is, as scientists, we should make it our business to seek out the observation that would falsify our law. Upon observing such a falsifier, we then discard that particular universal law and move on to another one. This, in short, is his technique of 'falsification'. In this way, science 'prunes' our claims of operational knowledge in the most efficient and logical way possible; finding one falsifier is an easier task than seeking out every affirmative case of a universal law. Indeed, as Popper asserts, the latter may be something we cannot even do:

> any conclusion drawn (from this process of induction) may always turn out to be false: no matter how many instances of white swans we may have observed, this does not justify the conclusion that all swans are white.[4]



If we are to accept this model, then the role of science is to show us how we are being inaccurate in our claims to knowledge about the world. That is, if I believe that 'all swans are white', and a Popperian shows me a black swan, then my explanation of the way the world is will stand to be improved by my discarding of this universal claim.

Science isn't in the business of being constitutive of our operational knowledge, it is to be understood as that which evaluates and refines it. What we commonly think of, as scientific claims to knowledge, are only tenuous hypotheses proposed with the intention of proving them wrong.

There are however, some strange consequences of this Popperian methodology. They become obvious when we ask what happens when one of our universal laws happens to be right. That is, what if I claim 'all ravens are black' and they *are* in fact all black? For Popper, we can never have this claim confirmed. Due to its very nature as a universal law it will always go beyond what we have the right to assert.

If we never see a non-black raven this does nothing to help our claim. If we only ever see black ravens, this does not help to strengthen our claim either. In fact, for Popper, the best we can hope for is that a given claim is corroborated at one instance in time, it can remain the best theory we have whilst it sits in the probationary wings, awaiting its refutation. In sum, a universal law can never be judged to be right - even if it is - as in *principle* we may always discover a falsifier. Intuitively, this just seems in direct conflict with our understanding of what knowledge is. That is, we want it to be something substantial and dependable, not something that is only useful to us in one restricted time and circumstance. When we set out to explain the subset of operational knowledge claims that we commonly know to be 'scientific knowledge' we did not intend to discard its status as knowledge altogether.

Nevertheless, if we maintain that our operational knowledge consists of sets of universal statements about the world then Popper's revelations are right to question the status of this as 'knowledge'. If Popper is right, then all we can ever know with logical certainty is that on one occasion we have proved a universal law to be wrong.

So where does this leave our question as to what scientific knowledge consists in? As it stands, we have two answers, depending on which methodology of science we choose to favour. According to inductivism, we have a type of operational knowledge which takes the form of universal scientific laws that we induce from observing repeated instances of a phenomenon. These laws seem to give the sense of certainty that we require from knowledge, that what we are claiming to know has some grounding above and beyond the individual case that we perceive here and now. However in light of Humean and Popperian thought, such certainty seems to be purely illusory. On the other hand, falsificationism tells us that scientific knowledge consists in telling us about what is not the case, it is the process of showing how and why these laws fall short of a correct explanation of the world. As a consequence however, we cannot claim to have certainty in any positive claim about the world. Prima facie, it seems that neither of these options is immediately attractive. We seem to either be allowed a body of scientific knowledge which is logically flawed, or we have a rather paradoxical type of knowledge which only consists of judgements of what is *not* the case. In neither case does our common sense view of scientific knowledge demonstrate any kind of justifiable reliability.

Arguably, what seems to be the problem is the pivotal role both methodologies give to the idea of a universal law. By taking opposing stances as to the possible truth of a universal law it seems inductivists fail to accommodate the possibility of falsifying instances, whereas falsificationists rule out the possibility of substantive knowledge of any universal law. Does it really have to be the case that we either know something to always be the case, throughout time and space, or, on the other hand, that we simply can only know for certain what we don't know?

Perhaps what we need to ask is whether our claims to knowledge need have this 'universal' character. Instead let us consider a statement such as 'I know the sun will probably rise tomorrow based on what I have experienced'. What is the difference between this and a universal law like 'the sun will always rise'? Well, in the former I acknowledge that my experience is playing a role in my judgement. That is, I clearly admit to using a reference class that is limited in scope both in terms of the number of observations it depends upon, and the time scale over which they were obtained. What we are accommodating for is the *possibility* of error. Whilst we may claim to know something on the basis of the information we have, this is not the same as claiming that the state of the world will never deviate from how the scientific community has hitherto descibed it. This would mean that our body of scientific knowledge is not claiming to present universal truths, but only suggests that we can accommodate our observations of the world so far under certain probabilistic claims.

Nevertheless there are objections to this type of suggestion. The first of which concerns the risk of invoking what Popper calls pseudo-science.[5] From a Popperian standpoint, universal laws are not only the targets of scientific attack, they are a necessary part of scientific practice. That is, the problem with non-universal or 'probable' statements like the one I proposed above is that they do not admit of conditions in which they can be falsified, and hence ultimately, be rejected. If a statement cannot be falsified then it simply isn't scientific, and is merely a piece of 'psuedo-science'. However, this objection is not as damaging as it first appears. We can still find a valuable role for falsification even when we disregard universal laws. When we say something has always been the case and therefore it is *probable* that it will continue to be that way, evidence showing cases both where this is realised *and* where it is rejected are of the utmost importance to us. The first type of evidence strengthens the probability of a claim to knowledge and the second type weakens it. The difference in relation to Popper's account is that falsifying evidence is used to develop theories and not simply as a means with which to categorically discard them.

From a philosophical standpoint perhaps the more compelling objection that we should consider regarding these probabilistic knowledge claims is the relationship between knowledge and truth. When we held in place that knowledge existed in the form of universal laws it was easy to imagine how a given piece

of knowledge was either true or false: it depended quite simply on whether the law was right or wrong. However, in terms of these 'probabilistic' scientific claims I have introduced, it is not immediately clear that we can claim to know with such certainty whether they admit of categorical truth or falsity. That is, it is hard to see when we would reject a law as false. Even if we find a falsifier, it is hard to judge when its occurrence is improbable enough to warrant rejection of that law.

In other words, it is arguably a consequence of abandoning the idea of a universal law that we no longer have a black and white picture of when we have truth. However, even if this is the case, this is not to say that we cannot have knowledge. Truth and knowledge are not the same thing. We certainly like to think of truth as something definitive or categorical, and it is for this reason that I suggest truth to provide something like a normative condition for scientific knowledge. It is the job of science to refine theories in the face of evidence until they thereby come closer to what we call truth. This suggestion seems favourable with regard to the phenomenological aspect of knowledge, as although we would never claim knowledge of something that we suspected to be false, we are still prepared to rectify or adjust our claims to knowledge in the face of new evidence. In fact, in reality we accept that knowledge is 'not a precise conception' and that 'it merges into "probable opinion."'[6]

Besides, it is a further point that being 'true' in itself is not the only support a piece of knowledge can have. As Russell stated:

> A body of individually probable opinions, if they are mutually coherent, become more probable than any one of them would be individually. It is in this way that many scientific hypotheses acquire their probability.[7]

Although this 'coherence' of knowledge claims with one another cannot guarantee that any of them are actually true, it still offers a kind of structural support. That is, although science may not provide definitively true knowledge, it does offer an internally supportive network of claims from which explanations and propositions can be extracted and put to use.

Arguably, science is in the business of working toward the truth, and if we had it already science would be redundant. I think we can accept that our claims to scientific knowledge are true only in the face of what has been experienced so far. Although scientific method may not produce infallible claims to knowledge, it must be borne in mind that it is only a *potential* for error that we admit. This does not equate to our being in a state of constant uncertainty. On the contrary, it seems that a degree of error is something we would be wise to admit. Or, as Russell put it:

> all our knowledge of truths is infected with *some* degree of doubt, and a theory which ignored this fact would be plainly wrong[8]

University of Warwick

Bibliography:

Hume, D. *A Treatise of Human Nature* (2001). Oxford University Press: Oxford.

Popper, K. (1959). *The Logic of Scientific Discovery.* Basic Books: New York.

Popper, K. (1963). *Conjectures and Refutations.* Routledge & Kegan Paul: London.

Russell, B. (1912) *The Problems of Philosophy.* (2001 edition) Oxford University Press: Oxford

[1] I should add that I don't wish to beg the question against Wittgensteinian concerns of whether behaviour is rule governed or rule guided, either term will do for my purposes here.

[2] Although the arguments against necessary connection of cause and effect and the problem of induction are often discussed separately, it suffices for our purposes to deal with them as striking the same target. For the original discussions see *Treatise of Human Nature* Book 1, Part 3, especially §6 & §14.

[3] See Russell (2001 edition) p.35

[4] See Popper (1959) p.27

[5] For a full presentation of Popper's views see Popper (1963) pp.33-39.

[6] Both of these quotes are from Russell (2001 edition) p. 78

[7] Ibid p. 81

[8] Ibid p. 78

Notes On Contributors

Jennifer Booth

is a philosopher at the University of Warwick, currently studying for her doctorate on self-conception, agency and elusiveness. She has taught on historical and contemporary epistemology and on the philosophy of science. Her current research interests are within the philosophy of mind and action.

Mat Carmody

started studying philosophy at Queens' College Cambridge before moving onto King's College London to an MPhil on holistic theories of meaning and a PhD on vagueness. He taught at King's College for a year (2001-2002) after completing his doctorate. He then took up a position in the Department of Logic in the Nicolaus Copernicus University in Torun, Poland. Since returning in the summer of 2003, he has been working in London and continuing his research into the sorites paradox. He is currently preparing a couple of papers outlining an anti-realist strategy for tackling it. Contact: mat.carmody@orange.net

Peter Goldie

is a lecturer at King's College London. His main philosophical interests are in ethics and the philosophy of mind, and increasingly in aesthetics. So far as his current research interests are concerned, he continues to have an enduring interest in emotion and character, and is becoming increasingly absorbed in considering the vast implications of the idea that everyday explanations are narrative in form. His publications include *The Emotions: A Philosophical Exploration* (Oxford: Oxford University Press, 2000), and most recently *On Personality* (London: Routledge, 2004). He has also edited a collection of essays on the emotions entitled *Understanding Emotion: Minds and Morals* (London: Ashgate, 2001).

Christopher Hamilton

is lecturer in philosophy of religion in the Department of Theology at King's College London. He is the author *Living Philosophy: Reflections on Life Meaning and Morality* (Edinburgh: Edinburgh University Press, 2001) and *Understanding Philosophy for AS Level* (Cheltenham: Nelson Thornes, 2003). He has also published numerous articles in academic journals, and has taught extensively at sixth form level as well as at university.

Alan Thomas

is head of the department of philosophy at the University of Kent at Canterbury, where he focuses on Moral, Political and Social Philosophy, Wittgenstein, and Kant's Philosophy of Mind. He is the author of Value and Context: *the Nature of Moral and Political Knowledge,* and *Thomas Nagel: a Study of his Philosophy.* He is also the editor of the *Cambridge Companion to Bernard Williams.* All three works are forthcoming.

Garrath Williams

is lecturer in philosophy at the University of Lancaster. He completed both his undergraduate and graduate studies at the University of Manchester, and has taught at the universities of Central Lancashire, Manchester, and St Andrews, as well as the European Academy in Bad Neuerahr-Ahrweiler. He specialises in moral philosophy, political thought and applied ethics. His numerous published articles include 'Nietzsche's Response to Kant's Morality' in *Philosophical Forum 30,* 'Normatively Demanding Creatures: Hobbes the Fall and Individual Responsibility' in *Res Publica 8,* and 'Blame and Responsibility' in *Ethical Theory and Moral Practice 6.*

Notes For Contributors

Content

We welcome articles on any area in philosophy. Papers may be broad or narrow in their focus (for instance a discussion of the mind/body problem, or an analysis of Hume's treatment of causation in the *Enquiry*). We would particularly encourage contributions which reflect original research on the following philosophical themes: epistemology, metaphysics, philosophy of religion, ethics, philosophy of mind, philosophy of science, political philosophy, religious ethics; and texts, such as: *The Republic, The Nicomachean Ethics, The Meditations, An Enquiry Concerning Human Understanding, Beyond Good and Evil, On Liberty, Existentialism and Humanism, The Problems of Philosophy, Language Truth and Logic*.

The articles should be around 3000-4000 words.

Style

The language used in the articles should be as non-technical as possible whilst preserving the richness of the arguments. Where technical terms are unavoidable they should be explained and examples offered.

Notes should be presented as endnotes. Textual references should be presented in the following format: Barry Stroud, *Hume* (London: Routledge, 1977), 77-91.

Presentation

Articles should be written in *Word* (any version).

Contributions

Articles for this journal are currently written by a panel of philosophers from a variety of universities in Britain, Australia and the United States, whose work is edited by the journal's editorial board. We invite unsolicited contributions from philosophers working in any field. The contributions should be submitted via email attachment to rjp@rutc.ac.uk

Copyright

The RJP retains the option of reprinting published articles in later RJP publications. Authors may republish articles with the journal's permission provided that they acknowledge that those articles were first printed in the RJP. Papers should only be submitted if the author is willing and able to be bound by the conditions set out in this paragraph.

Contributors

Richmond upon Thames College

is a large further education college located in
Twickenham, West London
offering 16-19 students one of the widest
curriculum choices in the country.

Last year, we came top of all
London further education colleges
in the Times league tables
and we are proud of our reputation
for achieving excellent results year after year.

We are well known nationally for
our high quality staff, excellent student support systems
and the inclusive education we offer to all our students.

If you would like to find out more about us -
please contact our Course Information Unit on

020 8607 8305 / 8314

or visit our website on

www.rutc.ac.uk

Richmond upon Thames College

[Subscribing]

to the *RJP*

The RJP comes out in Autumn, Spring and Summer. To subscribe you need to select the appropriate price from the table and complete the mailing information at the bottom of the form. All prices include post and packaging. The bottom half of the form should then be detached and sent with a cheque made payable to Richmond upon Thames College to the address below:

RJP Subscriptions
Philosophy Department
Richmond upon Thames College
Egerton Road
Twickenham
London TW2 7SJ
United Kingdom

Please allow one week for delivery in the UK, and two weeks for the rest of the world.

Annual Subscription : Current Rates

Institutional Subscriber in the UK	£33.00
Individual Subscriber in the UK	£18.00
Institutional Subscriber from the rest of the EU	63.00 Euros
Individual Subscriber from the rest of the EU	40.50 Euros
Institutional Subscriber outside the EU	$67.50 US
Individual Subscriber outside the EU	$45.00 US

Resubscription YES / NO From issue .....

[subscription]

Your name: ..... Your organisation's name (if appropriate) .....

Address .....

..... Post code/Zip code.....

Telephone number (including full international dialling code) .....

Email address .....

I enclose a cheque for the amount of ..... to purchase an annual subscription

Signed ..... Date .....

[Philosophy]

The Richmond Journal of **Philosophy**

Richmond Journal of Philosophy
Philosophy Department
Richmond upon Thames College
Egerton Road
Twickenham
Middlesex
TW2 7SJ
United Kingdom

Tel: 020 8607 8000
Fax: 020 8744 9738
Email: rjp@rutc.ac.uk
www.rutc.ac.uk/rjp

Richmond upon Thames College

ISSN: 1477-6480

Design & Production Marketing for Education 01282 612222 [ref 27353 - 07/04]